DOCUMENT RESUME

ED 057 217 VT 014 260

AUTHOR	Panitz, Adolf; Olivo, C. Thomas
TITLE	National Occupational Competency Testing Project. A Consortium for Occupational Competency Testing of Trade and Industrial/Technical Teachers. Phase I: Planning--Organizing--Pilot Testing. Volume 3, Handbook for Developing and Administering Occupational Competency Tests.
INSTITUTION	Rutgers, The State Univ., New Brunswick, N.J. Dept. of Vocational-Technical Education.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
PUB DATE	1 Feb 71
NOTE	171p.
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	Administrator Responsibility; Bibliographies; Educational History; Educational Philosophy; *Performance Tests; *Resource Materials; *Teacher Evaluation; *Test Construction; Test Results; Trade and Industrial Education; *Trade and Industrial Teachers

ABSTRACT

Much effort has been directed toward the evaluation of human characteristics and abilities. This handbook is an outgrowth of an extensive study of occupational competency testing and is intended to assist in the development, administration and evaluation of written and performance tests wherever occupational competency evaluation may be fundamental to employment, upgrading or promotion. This report: (1) provides the historical background and philosophical concepts of occupational competency testing, (2) outlines test development procedures, (3) discusses test administration, (4) considers the evaluation of occupational competency tests and test results, and (5) presents findings concerning national occupational competency testing. This is an outgrowth of the report on "The state of the art" and the experiences gained in pilot testing of two tests. List of resources, references and a bibliography complement this report. Related documents are available as ED 051 378, and VT 014 261 in this issue. (GEB)

ED 057217

NATIONAL OCCUPATIONAL COMPETENCY
TESTING PROJECT

Phase I: Planning – Organizing – Pilot Testing and Establishing the Feasibility of a National Consortium

Volume 3

HANDBOOK FOR DEVELOPING AND ADMINISTERING OCCUPATIONAL COMPETENCY TESTING

Dr. Adolf Panitz and Dr. C. Thomas Olivo

Research Project No. 8-0474 to
Department of Vocational-Technical Education
Graduate School of Education
Rutgers, The State University, New Brunswick, New Jersey

By U. S. Department of Health, Education and Welfare
Office of Education
National Center for Educational Research and Development

VT 014260

February 1, 1971

ED057217

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

NATIONAL OCCUPATIONAL COMPETENCY

TESTING PROJECT

A Consortium for Occupational Competency Testing of Trade and Industrial/Technical Teachers

Phase I: Organizing - Planning - Pilot Testing

Volume 3

HANDBOOK FOR DEVELOPING AND ADMINISTERING OCCUPATIONAL COMPETENCY TESTS

Dr. Adolf Panitz and Dr. C. Thomas Olivo

Research Project No. 8-0474 to
Department of Vocational-Technical Education
Graduate School of Education
Rutgers, The State University, New Brunswick, New Jersey

By U. S. Department of Health, Education and Welfare
Office of Education - Bureau of Research

February 1, 1971

The research reported herein was performed pursuant to a grant from the Office of Education, U. S. Department of Health, Education and Welfare. Contractors under Government sponsorship are encouraged to express freely their professional judgement in the conduct of their project. Points of view or opinions stated do not, therefore, necessarily represent the official Office of Education position or policy.

NATIONAL OCCUPATIONAL COMPETENCY TESTING PROJECT

PLANNING COMMITTEE

Dr. C. Thomas Olivo, Project Director
Dr. Adolf Panitz, Associate Director

PRINCIPAL INVESTIGATORS

Dr. Carl Schaefer, Chairman
Department of Vocational-Technical Education
Graduate School of Education
Rutgers, The State University
New Brunswick, New Jersey

Dr. Melvin Barlow, Director
Division of Vocational Education
Graduate School of Education
University of California
Los Angeles, California

Dr. Richard Nelson, Chief
Bureau of Industrial Education
California State Education Department
Sacramento, California

Dr. C. Thomas Olivo, Professor
Department of Industrial Education
Graduate College of Education
Temple University
Philadelphia, Pennsylvania

REPRESENTATIVE LEADERS

Mr. Walter Bialobrzeski (Connecticut)
Dr. Calvin J. Cotrell (Ohio-Research Center)
Dr. Edward Hankin (Georgia)
Dr. Ernest Kramer (Washington)
Dr. Gordon McMahon (New York)
Dr. Robert M. Reese (Ohio)
Dr. Byrl Shoemaker (Ohio)
Dr. Merle Strong (Wisconsin)

EX OFFICIO

Dr. Lance Hodes (Washington, D. C. - Project Officer)
Dr. Lawrence Braatan (Washington, D. C. - Resource Consultant)

P R E F A C E

Much effort has been directed toward the development of procedures and instruments for the evaluation of human characteristics and abilities. Frequently, for the purpose of predicting future achievement, tests have been developed to measure general achievement.

The literature on general achievement tests is rather extensive. However, the evaluation of occupational competence seems to attract attention only during periods of emergencies. World War I brought on the Oral Test, developed by the U. S. Army for the purpose of more effective classification of millions of civilians for Army occupations.

When the depression years created millions of unemployed, the United States Employment Service carried on extensive efforts to classify people according to certain skill levels. Oral tests were further developed. World War II brought additional refinements and a limited introduction of performance evaluation.

The Civil Service has played a major role in developing testing programs for the selection of individuals for a wide variety of jobs. Many of the Civil Service tests are of the written type. But, in recent years, practical performance tests have been introduced.

Similarly, the military, faced with an urgent need for better manpower utilization, has carried occupational competency testing to a fairly sophisticated level in selected and special occupational areas.

Vocational industrial-technical education, experiencing rapid expansion and the need for continuous qualitative improvements, is faced with the problem of selecting competent tradesmen and skilled technologists for teaching positions in the many varied vocational education programs. The need for valid and reliable measures for evaluating the competency of individuals from industry is urgent.

Despite the great need for such measurement, the effort to develop reliable instruments and techniques of measurement has received less attention than the need demands. Published information for test development is sparce indeed.

This "Handbook for Developing and Administering Occupational Competency Tests" is an outgrowth of extensive study of occupational competing testing on a national scale. It represents a summary of the best existing practices and procedures employed by the Civil Service, the military, various State Education Departments, labor, industry, vocational industrial-technical teacher education institutions, and the life-time experiences of the authors.

As the title implies, this "Handbook" is intended to assist in the development, administration and evaluation of written and performance tests wherever occupational competency evaluation may be fundamental to employment, upgrading or promotion.

The material has been organized in four main parts: Part I provides the historical background and philosophical concepts of occupational competency testing. These relate to the selection of tradesmen and technicians from industry for teaching in vocational programs; the nature of written, performance and oral tests, and their advantages and disadvantages.

Part II is concerned with test development procedures, the personnel involved in test construction and actual test construction. Examples are provided for illustrative purposes. The procedures involved in test administration are discussed in Part III. Part IV is devoted to the evaluation of occupational competency tests and test results. The findings concerning national occupational competency testing are reviewed in Part V.

Appendices of significant resources, references and a bibliography complement this report.

The contents of this Handbook, particularly the recommendations, reflect a careful synthesizing of the best-tested experiences interpreted against many, many years of practical work in occupational competency testing by the project director and codirector. The contents, therefore, do not necessarily reflect the position or policy of either Rutgers University or the United States Office of Education, and no official endorsement by either organization should be inferred.

New Brunswick, New Jersey
February 1, 1971

Adolf Panitz
C. Thomas Olivo

A C K N O W L E D G M E N T S

Phase I of the National Occupational Competency Project sought to establish the feasibility of pooling national resources into a Consortium of States for Occupational Competing Testing. Another objective related to the development and application of the most effective methodology and knowledge to the preparation, administration, pilot testing and reporting of occupational competency tests in major trade and industrial-technical occupations. Preliminary to this work was the comprehensive study on "The State of the Art".

This Handbook is a logical outgrowth of, (1) the report on "The State of the Art" of occupational competency testing which related to the efforts of people and organizations involved in developing measures for occupational competency, and (2) the experiences gained in pilot testing of two tests in four widely separated geographic states.

Throughout the project the lack of a practical handbook for the development and administration of occupational competency tests became crystal clear. Some institutions had prepared directions for test development and others for test administration. However, no comprehensive publication was available dealing with the most recent methodology and all aspects of test construction and administration.

Acknowledgment is made to those who assisted in this massive undertaking. First, recognition is made of the Principal Investigators whose perseverance achieved recognition of the critical need for occupational competency evaluation, and who were instrumental in obtaining financial and other professional support to carry on this project.

Credit is due Dr. Carl Schaefer for initiating the proposal and providing leadership from the preliminary stages to the funding and the successive stages of implementation.

Dr. Melvin Barlow, Dr. Richard Nelson, and the Project Director who, with Dr. Schaefer, represented the principal investigators, gave direction and active support. Recognition is also extended to the Planning Committee for direction, review of project materials, and other active assistance.

Among the individuals who contributed generously of their time and counsel, Dr. Gordon McMahon provided the project with materials resulting from many years of experience in occupational test development and his continuing leadership in this field in New York. Mr. James Peterman, Educational Advisor at the Great Lakes Naval Training Center, graciously made material on test item development available. Dr. R. O. Waldkoetter, Chief Evaluation and Analysis Branch, Enlisted Evaluation Center, Ft. Benjamin Harrison, Indianapolis was most helpful in providing material on competency test development and evaluation. Mr. Mangiaracina, Assistant Education Advisor, and Mr. C. Rudsinsky, Chief Test Developer, United States Army Signal Corps Center, Ft. Monmouth, New Jersey, were equally helpful in making competency performance test information available. Dr. Lance Hodes and Dr. Lawrence Braatten, of the United States Office of Education, assisted materially in providing resource materials and through a continuous review of project plans and progress.

Appreciation is, also, expressed to Dr. Benjamin Shimberg, of the Educational Testing Service, Princeton, New Jersey, and Dr. L. L. Wallace, of the Psychological Corporation, New York City, for their helpful cooperation on a number of problems relating to occupational competency testing.

A personal word of appreciation is expressed to Mrs. Mary Opalka for her patience and work in typing the manuscript for final reproduction.

Others who assisted in providing materials essential for the total study are identified in other appendices of resource personnel and services. Finally, recognition is made of all who participated in the initial efforts to lay the groundwork for a National Consortium of States for Occupational Competency Testing. Their deliberations and support represent a very vital force in ensuring that the results of the project are valuable to their needs and are possible of implementation.

F O R E W O R D

Instruction and learning in vocational education are bridged by the application of skills and information to a particular job. Every time a student has satisfactorily completed a job, he has, in fact, completed a performance test. Evaluating the finished product by checking the corrected trouble in a piece of service equipment, the function of a replaced part, or the tuned-up engine, are steps practically automatic to the skilled tradesmen and relatively simple to perform. The measurement is specific and measurable: ---does the part function properly ---is it made according to specification ---are all dimensions within prescribed tolerances?

It is more difficult to appraise such factors as: the correct procedure employed, the speed with which the job was done, and the manner in which tools, equipment and machines were handled. Although the finished product may meet all the specified requirements, there is no assurance that the learner fully understood the underlying principles involved and the technical knowledge required. While it is assumed that he did, there may be no measurable evidence. The teacher forms a judgement of the student's progress by asking questions, giving assignments in writing, and observing the learner in action. The type of work the student does indicates a level of manipulative proficiency and answers to questions provide a measure of understanding.

The informational factors, combined with the instructor's understanding of the requirements of the trade, enable him to evaluate the learners achievement and proficiency.

The key to effective instruction in vocational programs is the occupational competence of the instructor. How can such competency be evaluated? A number of procedures have been employed over the years and they consist basically of four elements:

1. appraisal of a teacher candidate's experience in the form of recommendations from employers,
2. oral questioning during an interview,
3. written tests concerning the knowledge, information and judgements of the occupation, and
4. actual performance under real or simulated working conditions.

Whatever limitations these procedures contain, they are a step in the direction of establishing some objective evidence of occupational competence over purely subjective judgements.

In recent years, there has grown an increasing concern about accurate and reliable testing procedures which can provide helpful information for evaluating a prospective teacher's knowledge and skill of the trade which he expects to teach, and the occupational proficiency which he has acquired.

This Handbook contains no magic formula for test planning, development, validation, nor is there a simple or inexpensive solution to the problem. Constructing a good test is a demanding, time-consuming task, challenging the best creative effort of a tradesman, test specialist and the test administrator. However, a summary of current best thinking and practices in occupational test construction and administration, presented in convenient form for reference, should prove useful to individuals responsible for the selection of tradesmen for teaching, for certification, and for admission to undergraduate and graduate programs of study.

Carl Schaefer

New Brunswick, New Jersey
February 1, 1971

CONTENTS

Page

Planning Committee ii

Preface iii

Acknowledgements v

Foreword vii

PART ONE: HISTORICAL BACKGROUND AND PHILOSOPHICAL CONCEPTS OF OCCUPATIONAL COMPETENCY TESTING

CHAPTER ONE: TYPES AND PURPOSES OF OCCUPATIONAL COMPETENCE TESTS . . 1

1. PURPOSE OF TESTING ON A NATIONAL SCALE. 1

A. Concern for Reliable and Valid Occupational Competency Tests. . . . . .
B. Constraints to Test Development. . . . . .
C. Dissatisfaction with Existing Instruments . . . .
D. Duplication of Effort . . . . . .
E. Economy and Efficiency of a National Effort. . . .

II. HISTORICAL BACKGROUND OF TESTING. 3

A. Quantitative Measurements. . . . . .
B. Units of Measurements . . . . . .
C. Measuring of Human Characteristics. . . . . .
D. Uncertainties of Performance Measurements . . . .
E. Written Tests are Dominant . . . . . .

III. BASIC CONCEPTS AND CHARACTERISTICS OF DIFFERENT TYPES OF OCCUPATIONAL COMPETENCY TESTS. 4

A. General Concepts. . . . . .
B. What is an Occupational Competency Test? . . . .
C. What are Job-Like Tasks? . . . . . .
D. Relationship of Performance Testing of Pencil and Paper Testing . . . . . .
E. Content of Written Proficiency Tests . . . . . .
F. Types and Uses of Performance Proficiency Tests . .
G. Specific Types of Tests . . . . . .
1. Recognition Test
2. Simulation Test
3. Work Sample Test
H. Uses of Performance Tests . . . . . .

Page

I. Oral Examinations for Appraisal of Occupational Competency . . . . . .
1. Historical Development and Use of Oral Examinations . . . . . .
2. Characteristics of Oral Tests . . . . . .
3. Limitations of Oral Tests. . . . . .
4. Preparation for the Oral Test . . . . . .
5. Oral Test Development Sequence . . . . . .
J. Performance Tests of Aptitude. . . . . .
K. Purposes Served by Occupational Competency Tests for Future Vocational Teachers . . . . . .
1. Admission to Trade and Industrial/Technical Teacher Education Programs. . . . . .
2. Competency Tests for Temporary or Permanent State Certification . . . . . .
3. Occupational Competency Tests for Advanced Standing in Collegiate Programs of Study Leading to the Baccalaureate Degree. . . .

CHAPTER TWO: ADVANTAGES AND DISADVANTAGES OF VARIOUS TYPES OF TESTS FOR OCCUPATIONAL COMPETENCY EVALUATION 18

I. LIMITING FACTORS IN TEST SELECTION 18

A. The Nature of the Trade or Occupation . . . . . .
B. The Type and Amount of Machinery and Equipment Required . . . . . .
C. The Materials, Parts and Supplies Needed . . . . . .
D. The Use to be Made of the Test Results. . . . . .
E. The Range of Skills and Information to be Evaluated .

II. ADVANTAGES AND DISADVANTAGES OF VARIOUS TESTS. 20

A. Advantages of Work-Sample Tests . . . . . .
B. Disadvantages of Work-Sample Tests . . . . . .
C. Advantages of Written Tests . . . . . .
D. Disadvantages of Written Tests . . . . . .

III. OCCUPATIONAL COMPETENCE MEASUREMENT WITH WRITTEN AND PERFORMANCE TESTS 22

A. Value of Multiple-Choice Test Items for Written Tests
B. Need for Both Written and Performance Tests

PART TWO: TEST DEVELOPMENT

CHAPTER THREE: TEST DEVELOPMENT PROCEDURES 23

I. THE IMPORTANCE OF PROPER PLANNING 23

Page

II. OUTLINE OF STEPS IN PLANNING WRITTEN TESTS 23

A. Sources of Related Job Classifications . . . . . .
B. Procedures for Determining Occupational Competencies to be Measured . . . . . .
C. Job Descriptions and Text Specifications . . . . .
D. Alternatives for Preparing a Job Description to Establish Test Content . . . . . .
E. Test Items, Test Construction and Evaluation . . .

III. PLANNING PERFORMANCE TESTS 31

A. General Considerations . . . . . .
B. Steps in Planning a Performance Test . . . . . .

IV. EVALUATING THE PERFORMANCE TEST 32

A. Work Method or Finished Product . . . . . .
B. Identifying Craftsmen Characteristics . . . . . .

V. PARTS OF THE OCCUPATION REQUIRING EVALUATION 33

A. Aspects that Represent Typical Jobs . . . . . .
B. Establishing Critical Items to be Measured . . . .

VI. TERMS DESCRIBING AREAS FOR SPECIFIC TEST CONTENT . . . 34

A. Variations in Test Terms . . . . . .
B. Terms Defined . . . . . .

VII. SUMMARY EXAMPLE: CONSIDERATIONS FOR THE DEVELOPMENT OF WRITTEN AND PERFORMANCE TESTS FOR ELECTRONICS INDUSTRIES OCCUPATIONS - COMMUNICATIONS 36

A. Test Purposes . . . . . .
B. Test Standards . . . . . .
C. Methodology . . . . . .
D. Approach to Occupational Competancy Test Development . . . . . .
E. Test Utilization . . . . . .
F. Test Scope - Electronics Industries Occupations - Communications . . . . . .
G. Job Description: Electronics Industries Occupations - Communications . . . . . .
H. Occupational Competency Measuring Instruments
1. Written Test . . . . . .
2. Performance Test . . . . . .
I. Scoring Methods . . . . . .
J. Rating Form . . . . . .

Page

CHAPTER FOUR: CONSTRUCTION OF TESTS 41

I. PERSONS INVOLVED IN TEST CONSTRUCTION 41

A. The Role of the Subject Specialist . . . . . .
B. The Role of the Vocational Teacher . . . . . .
C. The Role of a Committee of Specialists . . . . . .
1. Function of the Committee . . . . . .
2. Size of Committee. . . . . .
3. Committee, Direction and Working Procedures .

II. FUNCTION OF TEST SPECIFICATION 43

A. Range of Test Specifications . . . . . .

III. FORMULATION OF TEST ITEMS. 45

A. Choice of Test Items . . . . . .
B. Popularity of Multiple-Choice Test Items . . . . .
C. Changing Alternative Types of Test Items . . . . .
D. Summary of Objective-Type Tests . . . . . .

IV. VARIATIONS AND USE OF MULTIPLE-CHOICE TEST ITEMS . . . 50

A. The Stem as a Question or Incomplete Statement . .
B. Responses Requiring Selection of One Right Answer or the Best Answer. . . . . .
C. Types of Multiple-Choice Items. . . . . .
1. Test Items Relating to Definitions and Terms.
2. Test Items Relating to Purpose . . . . . .
3. Test Items Relating to Cause . . . . . .
4. Test Items Relating to Effect. . . . . .
5. Test Items Relating to Association . . . . .
6. Test Items Relating to Recognition of Error .
7. Test Items Relating to Identification of Error. . . . . .
8. Test Items Relating to Evaluation . . . . .
9. Test Items Relating to Differences . . . . .
10. Test Items Relating to Similarity . . . . .
11. Test Items Relating to Arrangement . . . . .
12. Test Items Relating to Incomplete Arrangements . . . . . .
13. Test Items Relating to a Common Principle. .
14. Test Items Relating to Controversial Subjects. . . . . .
15. Test Items Requiring the Most Inclusive Answer . . . . . .

V. TEST ITEMS IN NEGATIVE FORM 55

A. Test Items With Several Responses. . . . . .

VI. DEVELOPING MULTIPLE-CHOICE ITEMS 56

A. Rules for Formulating Test Items . . . . . .

Page

VII. CONTROL OF DIFFICULTY LEVEL FOR TEST ITEMS 62

A. Ways of Changing the Difficulty of a Test Item . .
B. Check Points for Evaluating Questions . . . . . .

VIII. RULES FOR DEVELOPING THE STEM OF THE TEST ITEM. . . . 63

A. Clear Central Problem. . . . . .
B. Pure Memory vs Understanding . . . . . .
C. Test Item Must Be Stated Briefly and Completely . .
D. Material Applicable to Solution . . . . . .
E. Applicable Correct Solution. . . . . .
F. Problem Stated in Positive Form . . . . . .
G. Avoid Making the Computational Problem Unintentionally Difficult . . . . . .
H. Include All Qualifying Information . . . . . .
I. When a Stem Contains Lengthy or Involved Qualifying Information, State Qualifying Information First and End Statement of Problem Requesting the Desired Information . . . . . .

IX. RULES FOR DEVELOPING THE ALTERNATIVES. 67

A. The Correct Response . . . . . .
B. The Distractors Representative of Errors . . . . .
C. Alternatives Should Be a Possible Answer . . . . .
D. Distractors Should Be Appropriate to Item Stem . .
E. Avoid Irrelevant Clues . . . . . .
F. Grammatical Sense . . . . . .
G. Overlap . . . . . .
H. Acceptable Alternatives of Responses. . . . . .

X. MISCELLANEOUS RULES AND POLICIES 74

A. Suggestions Pertaining to Numbers. . . . . .
B. Rules Pertaining to Capitalization . . . . . .
C. Punctuation After Alternatives. . . . . .

XI. VARIATIONS IN MULTIPLE-CHOICE TEST ITEMS. 76

A. Test Items Involving Multiple Completion . . . .
B. Application of Multiple Completion Test Items. . .
C. Appearance (Structure) of Multiple Completion Test Item Type . . . . . .
D. Hints in Construction of Multiple Completion Item .

XII. CLASSIFICATION ITEMS 79

A. Test Items with a Number of Key Topics . . . . .

Page

XIII. FIVE-RESPONSE MULTIPLE-CHOICE TEST ITEMS. 80

A. Test Type Most Widely Used in Objective Tests is Five-Response Multiple-Choice . . . . . .
B. Appropriateness of Fice-Response Multiple-Choice Completion Item Type . . . . . .
C. General Suggestions for Preparing Five-Response Multiple-Choice Items. . . . . .

XIV. GENERAL CHECKLIST FOR MULTIPLE-CHOICE ITEMS. 83

XV. LENGTH OF WRITTEN TESTS 84

XVI. PERFORMANCE TEST CONSTRUCTION 85

A. Skills Essential to Competency Evaluation . . . .
B. Process-Centered Work Sample . . . . . .
C. Product-Centered Work Sample . . . . . .

XVII. TIME LIMITS FOR WRITTEN AND PERFORMANCE TESTS 88

PART THREE: TEST ADMINISTRATION AND TEST RATING PROCEDURES

CHAPTER FIVE: TEST ADMINISTRATION. 90

I. GENERAL TESTING CONDITIONS 90

II. EXAMPLE OF DIRECTIONS FOR MACHINE INDUSTRIES OCCUPATIONS (MACHINE TRADES) CANDIDATES 91

A. General Information . . . . . .
B. Scope of the Written Examination . . . . . .
C. Performance Examination-General Information . . .
D. Scope of the Performance Examination. . . . . .

III. EXAMPLE OF DIRECTIONS FOR ELECTRONICS INDUSTRIES OCCUPATIONS-COMMUNICATIONS 96

A. General Information . . . . . .
B. Scope of the Written Examination . . . . . .
C. Performance Examination . . . . . .

IV. INSTRUCTIONS AND DIRECTIONS FOR EXAMINERS 101

A. Directions for Administering Written Examinations .
B. Directions for Administering Performance Examinations . . . . . .

Page

CHAPTER SIX: WRITTEN TEST RATINGS AND PERFORMANCE EVALUATION PROCEDURES 109

I. RECORDING TEST ANSWERS-WRITTEN TESTS 109

A. For Machine Cards . . . . . .
B. How Should Test Results Be Reported . . . . . .

II. RECORDING TEST RESULTS-PERFORMANCE TESTS. 110

A. Evaluation of Troubleshooting Items . . . . . .
B. Evaluations Including Time Factors . . . . . .
C. Evaluations Including Measurements of Accuracy . .
D. Evaluations of Finished Product . . . . . .
E. Evaluation of Proper Work Methods. . . . . .
F. Examiner Reliability . . . . . .

PART FOUR: TEST EVALUATION

CHAPTER SEVEN: EVALUATING OCCUPATIONSL COMPETENCY TESTS 118

I. KINDS OF VALIDITY 118

A. Content Validity . . . . . .
B. Concurrent Validity . . . . . .
C. Predictive Validity . . . . . .
D. Construct Validity. . . . . .

II. VALIDATION OF TESTS. 120

A. Logical Approach . . . . . .
B. Emperical Approach. . . . . .

III. ANALYZING PERFORMANCE TEST SCORES 121

A. Steps in Making the Analysis . . . . . .
B. Difficulty Level of the Test . . . . . .

IV. DETERMINING TEST RELIABILITY. 124

A. Factors that Determine Reliability . . . . . .
1. Test and Retest Method . . . . . .
2. Administration of Equivalent Forms . . . . . .
3. Split Halves Method-Subdivision of a Test Into Equal Parts . . . . . .

V. DIFFICULTY LEVEL OF TEST ITEMS 127

A. Difficulty Index . . . . . .
B. Difficulty Distribution . . . . . .

Page

VI. DISCRIMINATION OF TEST ITEMS 128

A. Discrimination Index . . . . . .

VII. LENGTH OF TESTS AND RELIABILITY 129

A. Reliability Coefficient and Length of Test . . .
B. Establishing Length of Tests . . . . . .

VIII. RANGE OF ABILITY AND TEST SCORES 130

IX. RANGE OF ABILITY REGARDED AS STANDARD 131

X. ITEM ANALYSIS OF TEST RESULTS 131

XI. ANALYSIS OF ALTERNATE CHOICES. 132

XII. TEST STANDARDS AND NORMS 134

A. Establishing Norms . . . . . .
B. Types of Norms . . . . . .

1. Computing Percentile Norms . . . . . .
2. Standard Scores. . . . . .
3. Normalized Standard Norms . . . . . .

PART FIVE: NATIONAL OCCUPATIONAL COMPETENCY TESTING PROGRAM

CHAPTER EIGHT: NATIONAL OCCUPATIONAL COMPETENCY TESTING PROGRAM: A NATIONAL URGENCY 140

I. THE NEED FOR A NATIONAL OCCUPATIONAL COMPETENCY TESTING PROGRAM 140

II. ADVANTAGE OF A NATIONAL TESTING PROGRAM 141

III. THE CONSORTIUM OF STATES AS A FACTOR IN RAISING STANDARDS AND RECOGNITION OF VOCATIONAL TEACHERS 141

PART SIX: APPENDICES OF RESOURCE MATERIAL 143

A. Selected References . . . . . .
B. Bibliography of Related Materials . . . . . .
C. Tables . . . . . .

I Sequence in Test Construction, Planning and Personnel Required Electronic Industries Occupations-Communications 26

II Selected Job Titles for Electronic Industries Occupations-Communications 27

Tables Page

III Job Clusters for Five Major Levels of Occupational Proficiency for Electronic Industries Occupations (Communications) 28

IV Job Descriptions (Occupational Competencies) for Electronics Occupations-Communications . . 29

V Test Item Planning for Industrial Electronics-Communication Occupational Competency Examination 30

VI Specification for Written Test - Electricity (Partial) 44

VII Test Plan Listing Objectives and Subject Areas . 44

VIII Multiple Level Outline 45

IX Summary of Common Objective-Type Test Item Forms and Characteristics 47-49

X Item Analysis of Results of a Carpentry Examination by Different Craftsmen 121

XI Frequency Distribution of Scores 122

XII Item Analysis of Test Results by Candidates Within Same Occupation 131

XIII Computation of Percentile Scores 136

Figures . . . . . .

1 Directions Summarized 78

2 Schematic AC-DC Receiver 86

3 Scoring Sheet 87

4 Sample Examination Report 107

5 Sample Report Form of Irregularities 108

6 Troubleshooting Evaluation Form 110

7 Time Standards for Job 111

8 Dimension Rating Form 112

9 Rating Form for Checking Sequence of Steps . . 113

10 Examiner's Performance Rating Scale. 115

Figures . . . . . Page

11	Examiner's Performance Evaluation Scale	116
12	Breakdown of Figure 11 (Back of Figure 8) . .	117
13	Curve of Test of Proper Difficulty Level for Maximum Discrimination	122
14	Curve of a Difficult Test	123
15	Curve of an Easy Test.	124
16	Analysis of Example 1	132
17	Analysis of Example 2	133
18	Analysis of Example 3	133
19	Analysis of Revised Example 3	134
20	Percentile Graph	137

Chapter I

TYPES AND PURPOSES OF OCCUPATIONAL COMPETENCY TESTS

I. PURPOSE OF TESTING ON A NATIONAL SCALE

A. Concern for Reliable and Valid Occupational Competency Tests

Vocational administrators and industrial teacher educators have been increasingly concerned for reliable measures to evaluate the occupational competence of prospective trade and industrial-technical teachers.

A number of states--notably, New York, California, Pennsylvania and Florida--have had the successful passing of occupational competency examinations a requirement for certification. Other states have started, or are planning to start, using such examinations. A number of Industrial Teacher Education Institutions are using such tests for admission to and advanced standing in undergraduate and graduate programs.

Many state vocational education officials have stated that their departments are not equipped to develop new examinations or to keep current ones up to date. There is general agreement that the development, printing, distribution, administration, scoring and validation impose a heavy burden upon limited resources. No state reported it was satisfied with its occupational competency testing program. Further, even the most highly industrialized states indicated their limitations to satisfactorily carry on such a program alone.

School administrators, when they employ a new trade or industrial-technical teacher, want to be assured that the candidate possesses the personal qualities which are necessary for success as a teacher. They are, also, concerned with the individual's professional preparation. But, these two essential prerequisites are of little value without evidence of competence in a trade or industrial-technical occupation which the individual is expected to teach.

The level of competence of the tradesmen is the keystone upon which the whole instructional program is built. Thus, the administrator (and the industry whom the program serves) must have assurances that the range of occupational competenc of the instructor is adequate to meet the full spectrum of youth and adult training needs.

B. Constraints to Test Development

Few states have been able to develop or participate in establishing accurate and reliable instruments suitable for evaluating the occupational competence of prospective trade and industrial-technical education teachers. These instruments must evaluate skills, related technology,

judgements and other understandings, mastered and further developed under actual wage-earning conditions. The development, administration and validation of occupational competency tests require human and material resources and finances that are in short supply even in leading industrial states. While each state recognizes that instructional excellence must be rooted in the occupational competence of each teacher, accurate measures for evaluating such excellence have not been developed to any practical degree.

C. Dissatisfaction with Existing Instruments

The need to establish a national coordinated effort for competency appraisal has been sounded in loud, continuous and clear voices, decade after decade. Then, in 1959, a survey by Dr. Carl Schaefer showed that sixteen states were using tests to evaluate trade competency. However, discussions with officials responsible for vocational education and teacher certification in these states revealed that considerable dissatisfaction existed with the quality of these tests.

In 1966, representatives of twenty-three states met at Rutgers University in a seminar to study the feasibility of providing trade competency examinations on a national base. The outcome of the seminar indicated that the development of occupational competency examinations nationwide would be a more efficient use of personnel and should provide a higher quality examination. Almost unanimous agreement was expressed that such examinations would be used by most of the states.

D. Duplication of Effort

In a recent survey of the state-of-the-art of occupational competency testing, it was established that twenty centers, located in sixteen states, are engaged in the preparation and administration of competency tests It was also found that twenty centers produced tests in the same trade area; fifteen centers duplicated their efforts for another trade area, and ten centers created tests for a number of additional trades. Each duplicated the efforts of the other. The duplication ranged from three to twenty tests in occupations little affected by special local or geographic conditions.

These tests were largely of the paper and pencil variety, although some require actual performance in simulated job situations. There is little evidence to indicate that test items were pretested or that the instruments were validated according to acceptable test development procedures. Little attention seems to have been given to such important technical considerations as reliability of scores or the objectivity of scoring procedures.

The survey furthermore revealed that most of the test development and administration is carried on by individuals with many other duties who are able to devote only a limited amount of time to this effort. At the present time, only three states have assigned one person each to a testing program. Many states expressed a need for help with the prepara-

tion and administration of trade proficiency tests for a teacher testing program.

E. Economy and Efficiency of a National Effort

A national effort would reduce or eliminate much costly duplication now found among a number of states. Substantial economies should be achieved and the quality of the tests greatly improved. National effort, resulting in accurate and objective evidence of occupational competency, would provide flexibility in the use of such tests; all states would have access to each examination. Each state would retain wide latitude in the use and interpretation of the results.

A national effort can meet the states' concern about the reliability of occupational tests and their own limited resources to establish the necessary evidence for occupational competency. A National Occupational Competency Testing Program would have a wholesome effect upon the vocational education teaching profession in a manner similar to the national tests for professional architects, accountants, and others.

II. HISTORICAL BACKGROUND OF TESTING

A. Quantitative Measurements

"Whatever exists, exists in some amount" has served as a slogan for a vast amount of effort that has gone into the construction of measuring instruments for teachers who wish to make an accurate report of the progress of their students.[1] When it is further asserted that "What exists, exists in some amount and can be measured" the slogan becomes a goal toward which test constructors are striving.

Measuring the amount of things is foundational to the worldwide scientific revolution of modern times. Precise measurement is one of the crucial elements of the scientific method of inquiry and work. The scientist observes and collects his facts with the aid of instruments of observation and measurement, organizes his facts in quantitative frequencies, treats these measured facts with mathematical methods, qualifies his findings with statistical reliability and draws conclusions only from such validated data.

B. Units of Measurements

To measure what can be measured requires "scales of equal units" so that every aspect of the physical universe can be measured with precision: weight, time, temperature, circular measure, electrical power, etc. The natural sciences became actualities as measuring instruments with scales of equal units perfected. In the building of these sciences, it was assumed that these were true facts which man could approximate by improving measuring instruments, methods of using the instruments, and the treatment of the measurements. The measuring movement in all the sciences devoted much energy to "cutting down error" in the observed facts.[2]

C. Measuring of Human Characteristics

Attempts at measuring human characteristics and school achievement date back well over one hundred years. Much early work was done in Western Europe, Germany and France, and later by Americans. Oscar Boros published his first Mental Measurements Yearbook in 1938. He listed 1181 instruments of mental measurement scales, tests, questionnaires, personal history records, rating scales, score cards and other objective instruments which were available that year, and 814 books, monographs and bulletins dealing with the movement. The number has substantially increased and descriptions have been published in subsequent yearbooks.

D. Uncertainties of Performance Measurements

In spite of the enormous amount of careful work measuring human characteristics and school achievement, measuring the performance of human beings still involves uncertainties. Much of the effort has been concentrated on acts of habit formulation in school skills: handwriting, arithmetic, linguistic, historical, mathematical and scientific facts and knowledge. Acts of thinking, reasoning and problem solving in the natural and social sciences; acts of creativeness, expression and appreciation applied to the work of the schools, are included in this massive testing effort.

E. Written tests are Dominant

Among the many forms of tests developed over the years, the written test dominates. The most common objective tests involve a number of different forms of these types of test items: multiple choice, completion, true-false questions, simple recall, matching, analogy, short essay, etc. The dominance of these type of test items is not surprising for the evaluation of student achievement has been the major concern of psychologists and educators alike. Paper and pencil tests, and test results, lend themselves to more readily objective treatment; the data can be treated more efficiently and acceptable evidence established.

While progress in the measurement of verbal skills, facts, and information; as well as, intelligence, the appraisal of personality traits, occupational interests, mechanical comprehension, spatial visionalization and abstract reasoning has been substantial, attempts at evaluating occupational skills and know-how by using written tests and performance tests have been relatively limited.

III. BASIC CONCEPTS AND CHARACTERISTICS OF DIFFERENT TYPES OF OCCUPATIONAL COMPETENCY TESTS

A. General Concepts

Guidelines for the development of occupational competency tests and their ultimate use as measuring instruments are not available. This Handbook contains material gathered from many sources not easily acces-

sible through the general media of publications. The guidelines are presented as major criteria which must be applied to the development of occupational competency tests. The practicality of developing written and performance tests for a given occupation can be determined through a review of the criteria. These guidelines and criteria have been prepared primarily for the purpose of relating competency proficiency evaluation to the selection of people from industry for teaching in vocational education programs.

B. What is an Occupational Competency Test?

An occupational competency test requires an individual to perform job-like tasks required of a craftsman or skilled technician within an occupation according to industrial standards of speed, accuracy, procedures, etc. The test is administered under well-controlled conditions to measure the individual's job proficiency through his ability to perform actual or job-like tasks. A well-defined goal is essential. The goal is the result the examinee is trying to achieve as he performs in a test situation. The goal should be the same for everyone and examinees must try to achieve the same results.

If one individual works for speed and the other for accuracy, the test does not measure what it should. The test must be conducted so as to give each examinee an equal chance to demonstrate his skill and occupational knowledge and to assure that the test conditions are the same for all examinees.

C. What are Job-Like Tasks?

Job-like tasks include those tasks, derived from an occupational analysis, which are representative of the on-the-job work performed by a competent individual in a particular occupation. For most occupations, these involve manipulative, as well as informational skills. The relationship between manipulative and knowledge skills varies from occupation to occupation, and requires careful analysis. It is important to test for both the manipulative proficiency and the technical knowledge, trade information, judgements, etc. essential to carrying out the work of the occupation at an acceptable competency level.

D. Relationship of Performance Testing to Pencil and Paper Testing

1. Performance tests are, generally, assumed to be superior to written tests in measuring the job behavior of tradesmen and technicians. However, they are time-consuming, expensive to develop and administer, and pose difficulties in scoring. While written tests are sometimes used in place of performance tests, they do not measure manipulative/performance competency.

2. For the measurement of occupational proficiency, it is essential to administer both performance and written tests. Some test specialists maintain that only performance tests are a measure of real occupational

proficiency. Others insist that some pencil and paper tests can be made job-like and are, also, performance tests. In occupational competency testing, neither of these points of view is accurate and applicable. While performance tests, in general, are more likely to be job-like than pencil and paper tests, certain aspects of information, facts, computations and interpretations can be more effectively measured through written tests. The determining factor, as to which test or combination of tests to use, is essentially "What is to be measured".

3. Tests can have the appearance of being good tests, even when they are not useful for measurement purpose. The original intent of the test constructor is nearly always good. However, the final test which he produces may or may not be adequate from the measurement standpoint, depending on what sacrifices and compromises or judgements are made in the course of the test development.

4. If practical considerations, such as, time, availability of machines or equipment, etc., have a dominant influence in determining what the performance tests are like, all of the original purposes may not be represented in the final product.

When time is a factor, the test constructor may provide directions on how to proceed which would enable the examinee to proceed faster. While this is permissible, it must not eliminate or reduce the opportunity to measure the intended skill. One of the major purposes of a test is to determine whether the examinee knows how to perform the task without being told how. If the examinee is told how to proceed, the test will only measure whether he can perform the task after he has been given much orientation. The test may then no longer measure what was originally intended. The performance test should go beyond pencil and paper tests and should measure the actual doing.

E. Content of Written Proficiency Tests

Written tests are usually easier to administer, quicker to score and provide more objective results. They are effective instruments to appraise certain aspects of occupational competence. The content of written tests should include the following areas of occupational competence:

1. The factual knowledge and technological (trade theory) information of the occupation.

2. Knowledge and understanding of the machines, tools, instruments, and apparatus of the occupation.

3. The properties and characteristics of materials and the parts, or systems, which are diagnosed, processed and replaced.

4. The necessary calculations required for setting up, adjusting, testing and operating the machinery, apparatus, test equipment used in the occupation.

5. The reading and interpretation of specifications, technical drawings, wiring and circuit diagrams, and other graphic representations.

6. The interpretation of scientific principles as they apply to the design and/or operation of machines, apparatus, instruments and the processing of materials, component parts and accessories.

7. Judgements and procedures involved in planning the work of the occupation, selection of materials and/or the replacement of component parts.

8. Classification of materials, component parts and specific tools, instruments and supplies.

9. General knowledge and work habits of the occupation, as well as specific requirements relating to occupational safety and hygiene, government regulations, and other conditions surrounding the occupation.

F. Types and Uses of Performance Proficiency Tests

A performance test for occupational proficiency generally involves a work sample test which requires the examinee to demonstrate his acquired skill by doing an actual segment of work using tools, materials, machines and equipment characteristic of the occupation for which the test is designed.

Although, for a long period of time, it was assumed by some that written tests provided valid information to establish a level of occupational competence, recent observations and the results of studies have indicated that written tests, alone, do not provide adequate and reliable evidence for this purpose. The relationship between the ability to answer test items on an examination and to perform in an actual job situation is very low.[3] The relation between verbal and mechanical intelligence has been shown by repeated use of tests to be very slight.[4] Therefore, it is essential that an applicant for teaching in trade and industrial-technical education demonstrate his job skill and the required job knowledge in a practical situation typical of the occupation.

1. A good performance proficiency test:

   a. Involves the measurement of critical skills of the occupation which distinquish the competent craftsman from the less competent individual.

   b. Includes selected manipulative processes which avoid extended repetition.

2. Well-designed performance tests must include:

   a. The actual operation of machines, apparatus, instruments and tools as utilized in a job situation.

   b. Step-by-step procedures in designing, forming, shaping, turning, cutting materials into a finished product, or diagnosing and assembling units and components.

   c. Step-by-step procedures in locating difficulties in equipment, instruments and/or apparatus.

   d. The planning and actual replacement of defective parts or components.

   e. The adjusting, maintaining, and utilizing of instruments of various kinds in carrying on the work of the occupation.

G Specific Types of Tests

There are three broad categories of tests which involve active "doing" on the part of the individual being tested:

- The Recognition Test
- The Simulated Condition Test
- The Work Sample Test

1. Recognition Test

The "Recognition Test" or "Identification Test", as the names imply, attempt to measure the individual's competence to recognize the essential characteristics of an operation, or the product of performance, or to identify objects by specific name, functional or operational value, etc. For example, an individual may be shown a series of splices of electrical wires and be asked to determine those that are correctly spliced and those that are incorrectly spliced. In other performance tests of the recognition type, the individual may be asked to identify parts of an automotive engine and their function. In order to achieve a maximum score, a testee would have to identify the engine parts in a definite order.

Value of Recognition Test

a. A performance test, with recognition items, is relatively easy to construct and can be adopted to a fairly wide range of subjects or operations. Although the test items measure some aspects of performance, they do not measure directly the individual's skill, technique or procedure.

The recognition test:

b. Provides a variety of activities suitable for testing candidates in both the general area of the occupation and specialized divisions,

c. Can be conducted in a well-equipped school shop laboratory or commercial establishment,

d. Provides occupational activities which the average instructor should be able to perform to a marked degree of proficiency,

e. Must include a variety of assignments in order to determine the all-around ability of the candidate,

f. Should provide assignments of different lengths of time for completion so that a variety of occupational practices may be completed within the time limit established,

g. Must include assignments which can be evaluated objectively by a competent and unbiased examiner. Its ratings must be objective, must differentiate and must distinguish between individuals who differ by a certain amount in trade ability. "The performance test signifies what is commonly meant by a man's competence to follow his trade, occupation or profession. Any complex set of coordinations which are required in a definite order and which characterize all men skilled in a given trade, thereby segregating them as a homogeneous group" [5] ...is a further definition of occupational competence.

2. Simulation Test

The simulation technique of imitating actual conditions was developed during World War II and is best represented by the link trainer which aided the Air Force in the training of pilots. Simulation tests are designed to isolate and duplicate the essential activities of a task, work, job, or sequence of operations which are necessary to analyze and correct a malfunction of a piece of equipment. The simulation technique has been carried to a very high degree of proficiency in the training program of astronauts.

In simulation tests, the candidates carry out realistic tasks in a setting which simulates a real situation. Simulation techniques have the advantage of economy of teaching/testing time, convenience, ease of administration and observation. For example, in the interest of safety, tests are often devised to demonstrate the condition under which a fatal electric shock will occur if the man on the job performs an operation incorrectly. Thus, in simulating work with high power lines, high voltage equipment in electronica or pneumatic clamping devices, a red light or a bell would indicate that an error has been made which

might well cause serious injury or fatality if the operation had been performed in this manner on an actual line.

Although the use of simulated conditions and miniature tests have certain advantages in the safe measurement of performances, test constructors must be cautions in assuming the degree of relationship between simulated or miniature tests and the more complete performance with actual equipment and under realistic conditions. What may seem to be a valid sample of an isolated task or a series of steps, may not be a valid sample of the total job situation. Simulated conditions of test performance, therefore, must be recognized and overcome if they are to serve the purposes intended.

3. Work Sample Test

A work sample test requires that the individual demonstrate his skill and work knowledge by doing a series of tasks, such as correcting difficulties in a piece of electronic equipment or completing a piece of work under actual work conditions. This type of test is realistic in that it requires the individual to carry through the actual operations that the job demands under normal working conditions.

In practical use, the performance test may consist of selected samples of a job, a product or part to be completed within a certain period of time; or the partial completion of an assignment or product. For many types of occupations, a completed job is not economical. For example, a complete engine overhaul in auto trades would be too time consuming. Generally, the more limited the sample job (which will predict the whole job) the greater the advantage as far as economical test administration is concerned. Therefore, it is important to find selected samples of performance that are sufficiently indicative of the individual's occupational competence to insure adequate measurement of his proficiency. [6]

Work sample tests have greater real-life appearances or face validity than any other type of test. When work sample tests are administered under standard conditions with standardized scoring procedures which have been carefully developed, they can provide valid and reliable estimates of job performance.

H. Uses of Performance Tests

Performance tests can, generally, be categorized according to one of five types of uses:

1. To evaluate proficiency and predict successful job performance.

2. To diagnose deficiencies in an individual's job performance,

3. To obtain a quantified estimate of job proficiency as of a certain point of time,

4. To provide a criterion for various uses or applications,

5. To evaluate competency at various levels of the occupation and, thus, serve as a means for selecting the most competent individual.

I. Oral Examinations for Appraisal of Occupational Competency

Industry, business and, to some degree, trades and industrial teacher education institutions utilize the interview as another means for appraising individuals for employment or for admission to teacher education programs. Conducted by competent individuals familiar with the techniques of interviewing, the interview represents another device for evaluating many subjective characteristics. As a measuring device, the interview must be conducted as a somewhat informal oral examination. At best, such an examination, even when carefully prepared, contains certain weaknesses. Oral examinations may provide: "The information which the man has with regard to certain elements of his trade. Only insofar as information is the reflection of experience in the trade, obtained at firsthand, will the question method give any indication of skill." [7]

1. Historical Development and Use of Oral Examinations

The Oral test was extensively used by the United States Army during World War I to assign soldiers to military jobs as closely related to their civilian occupation as was then possible. To separate the competent carpenter, machinist or electrician from those who said they were competent in these occupational areas, well-designed questions probably served their purpose.

During the depression years, the United States Employment Service faced the monumental task of classifying huge numbers of applicants according to occupational competency levels. A great deal of effort was devoted to the development of oral questions for over 150 trades and occupations to aid employment interviewers in determining more logically the competency levels of unemployed tradesmen and industrial workers.[8] In the late forties and early fifties, the Employment Services in the various states used the questions for job placement and identification of training needs. In recent years, the oral examination as the only means of selection has been discontinued in nearly all government employment offices.

Perhaps, this was the result of a widening belief, supported by recent research results, that there is a low relationship between knowledge of an occupation and the actual skill an individual must possess to carry out the more complex tasks of the trade.

However, limited numbers with a contrary view maintain that it is possible to separate the competent from the less competent individual by means of oral questions.[9] This group assumes that the larger percentage

of workmen who can perform the work of the occupation at an acceptable level of proficiency must be able to talk about their work in trade terms, provided the interviewer, or examiner, and the examinee speak the same occupational language.

2. The Characteristics of Oral Tests

In the appraisal of occupational competence one does not need to subscribe to either extreme view. However, it must be recognized that there are circumstances and conditions when neither written nor performance tests can be administered. Under those circumstances, an interview may well be the only means available to gain an overall impression on an individual's competence to express himself on matters of his occupation. Since this is one of the characteristics essential for teaching, the oral examination has value in roughly screening candidates.

As in all testing, careful effort must be made to: develop valid and reliable questions, establish a procedure for recording correct and incorrect responses, and to analyze the results when a group of individuals must be interviewed. Provision must also be made to secure the oral questions so that examinees do not exchange oral testing information during or after the interview.

In addition to the usual guidelines for preparing good test items, the following are emphasized as they relate to oral test items which apply to occupational experience and competencies:

a. The questions must be skillfully formulated and worded and must be concerned with familiar processes, materials, machines and equipment.

b. The questions must be clear so that they cannot be misunderstood. Suggested elements of correct answers should be recorded for use by the examiners.

c. The terms of the questions should be phrased in the vocabulary of the shop or occupation, and must reflect the level of difficulty experienced on the job.

d. The questions should avoid trade terms that are local or geographic.

e. The oral test items should represent all levels of occupational difficulty to differentiate among levels and distinguish capability.

f. The questions should aim toward a balance of test items representative of each major component part of a trade, as determined through an occupational analysis.

The United States Employment Service used oral tests almost exclusively during World War II to measure occupational competency. This governmental service now has standardized sets of oral tests covering over 126 skilled and semi-skilled occupations with instructional manuals for their administration and scoring.

Stead and Shartle, who were assisted in this effort to develop oral tests to measure occupational competence, reported eight basic kinds of questions:[10]

a. Questions dealing with <u>definitions</u> and calling for short descriptive answers.

b. Questions dealing with <u>modifications</u> and <u>limitations</u> of tools, materials and machines used in the trade.

c. Questions dealing with <u>the use of</u> tools, materials and machines of the trade.

d. Questions dealing with <u>procedures</u> by which the worker goes about doing the operations of the trade.

e. Questions dealing with the <u>location of things used</u>, <u>processed</u> and assembled in the trade.

f. Questions dealing with <u>proper names</u> of tools, materials, parts and processes.

g. Questions dealing with the purpose of <u>doing</u> things, <u>using</u> things, <u>locating</u> things, and the like, in the trade practice.

h. Questions dealing with <u>numbers</u> as related to tools, parts of machines, materials and the like.

3. <u>Limitations of Oral Tests</u>

If no other means of evaluating occupational competence is available, the oral test serves a useful, limited purpose. The following limitations must be considered:

a. The answers to oral test items are extremely difficult to treat statistically.

b. While multiple answers may show that a workman understands the whole process, it is rarely possible to give objective ratings to the answers.

c. There is the possibility that a single key element of a process will give an indication of knowledge or ignorance of the whole process.

4. Preparation for the Oral Test

The questions for the oral interview should be developed with the following in mind:

a. Is the test item applicable to the occupation?

b. Does the answer to the test item reveal or involve good occupational practice?

c. Should the test item be revised, modified or supplemented to yield suitable answers, or is it better eliminated?

d. Does the test item differentiate between levels of competency and distinguish the individuals having the specified competencies?

5. Oral Test Development Sequence

Oral test items should be developed according to accepted test development practices with emphasis on the following:

a. Prepare a test specification in which the major occupational areas included in the test are stated.

b. Test items should be compiled by occupationally competent test specialists and be adaptable to statistical treatment.

c. Oral test items should be written, typed or printed so that the examiner may refer to them in an unobtrusive manner.

d. The test items should be reviewed for internal validity, tested on a sample population, revised, etc.

e. The test items should then be standardized.

f. Oral test results should be treated statistically wherever possible.

g. The final items for an oral test should be selected and the test calibrated.

J. Performance Tests of Aptitude

In a performance test, the examinee is directed to carry out some manipulative activity representative of tasks performed in an occupation. Like written tests, performance tests can be classified as either aptitude or achievement tests according to the purpose for which the test is designed. The general form of an aptitude test is not much different from that of a proficiency test. Its major difference is in the relationship between the functions measured by the test and the requirements of the job for which performance is predicted.

Performance tests of aptitude are used before training starts. The results are intended to measure the individual's potential skill. The test function is to determine the capability of the individual to develop skills through training or the levels of such skill that he can develop, and, to a degree, how quickly he may reach a determined occupational goal.

Performance tests for aptitude consist of some task or group of tasks which, although not taken from the job directly for which the test is developed, nevertheless, predict success by measuring the aptitude suited to the needs of the job.

Examples of aptitude performance tests are numerous. They have been designed to test the individual's ability to assemble small parts, handle and place in special location pins with tweezers, fit uniquely shaped blocks together--others test the strength of the grip of hands, steadiness of hand, etc. These tests serve, primarily, to determine an individual's dexterity, his mechanical comprehension, mechanical reasoning, and visual acuity among many others. The value of performance tests of aptitude is, generally, limited to appraising broad categories of aptitude, such as mechanical ability or physical-motor abilities.

Such tests have been carefully developed and show the relationship between resulting scores and the known aptitudes of the subjects studied before the tests have been applied. These tests, since they do not test job knowledge, are suitable in selecting individuals with particular aptitudes for training for a particular payroll job.

K. Purposes Served by Occupational Competency Tests for Future Vocational Teachers

1. Admission to Trade and Industrial/Technical Teacher Education Programs

The teaching profession involves abilities, skills and know-how all its own. While there are some individuals who seem to be born with the gifts of the good teacher, the larger number must acquire the methods, skills and understanding. Experienced industrial teacher educators have observed that the experienced tradesman or skilled technician brings to the profession of teaching (besides his occupational skills and technical know-how) an attitude and outlook on life frequently very different from that of the purely college-trained teacher. They have, also, noted that work habits, relationships among people in an industrialized society, and a sense of practical responsibility acquired in the world of work carry

over into the experiences required by an individual to prepare himself for teaching his occupation.

While the industrial teacher educator is concerned with the personal qualities which are necessary for success in teaching and guides the individual through a program of professional studies, he, also, seeks to establish evidence of the candidate's level of competence in a trade or industrial-technical occupation. Towards this end, the applicant's record is evaluated. His prior work experience and recommendations of former employers are reviewed; sometimes by a committee of teachers and men from industry.

Although these procedures have worked, much of the information obtained depends on judgement and general impression. In recent years, there has been growing concern about the need for more reliable and valid instruments to establish the level of occupational competence through written and performance tests.

Two points need to be reemphasized at this time: (1) Occupational competence in the trades and industries and industrial/technical occupations is developed on the job according to a variety of training plans and programs; (2) Occupational competency represents the subject content competency for a course, curriculum or program.

Well-designed tests, properly validated and with their reliability thoroughly established, provide measures that are fair to individual applicants, provide a means of comparing the demonstrated achievement of applicants with accepted levels of proficiency, and eventually provide the necessary information for the development of practical acceptance norms an applicant will have to meet. Tests of this kind provide helpful information which the industrial teacher trainer can apply to correct occupational weaknesses in an otherwise qualified candidate, and to accept occupationally competent persons into an industrial teacher education program.

2. Competency Tests for Temporary or Permanent State Certification

Nearly every state recognizes that instructional excellence must be rooted in the occupational competence of the teacher. The State Plans of at least sixteen states have requirements involving some form of competency evaluation of trade and industrial/technical education teachers through testing. Nearly all states have established minimum experience requirements as one of the conditions for temporary or permanent certification. State officials concerned with vocational education have recognized the need for more accurate and effective means for establishing occupational competence, and are requiring or encouraging occupational competency tests for temporary or permanent certification. They have recognized that reliable test results provide objective evidence of subject content capability for certification and will lead to higher standards of teaching and learning achievement. Valid occupational competency testing will, also, improve the professional standing of vocational teachers and the whole teaching profession.

3. Occupational Competency Tests for Advanced Standing in Collegiate Programs of Study Leading to the Baccalaureate Degree

Granting college credits for employment experience is not new. It has been a slow and arduous process for colleges to accept the philosophy of a flexible system of credit examination. The ideal system of credit examination would assess and certify accomplishment on the basis of present performance, irrespective of the route that the individual traveled to achieve such competence.

Competency in an occupation is considered to include the skills, knowledges and judgements possessed by a typical journeyman. Competency also connotes human traits and capabilities in interpersonal and civic relationships...work habits, attitudes, team efforts, etc.

A national system of occupational competency examinations will permit individuals to participate in higher education with advanced standing. This recognized advanced standing will remove some of the barriers which now make it extremely difficult for vocational teachers to pursue advanced studies and achieve their degrees.

A national testing program will provide the evidence necessary to establish standards for the granting of advanced standing and encourage qualified individuals to strive for further professional improvement. It will enable the institutions to apply the results of their own requirements. Ultimately, the national occupational competency testing program cannot help but improve the professional standing of the vocational teacher. To digress, it may be stated, parenthetically, that comparable content capability measures, developed for all subjects and related areas of the total profession, would improve educational and training excellence in general, as well as specialized vocational education.

Importantly, it is here recorded that college credit granted for validated competency (validated through practical performance and written theory tests) should be equivalent to the credits for college work which leads to teaching in the academic areas--the fine arts, the liberal arts, and the practical arts. This not withstanding the fact that the greater number of practical skills and the related technical knowledge in the trades, industry and industrial technical occupations have been developed under actual employment conditions in industry.

Chapter II

ADVANTAGES AND DISADVANTAGES OF VARIOUS TYPES OF TESTS FOR OCCUPATIONAL COMPETENCY EVALUATION

The test constructor is faced with a number of limitations and obstacles in the search for the most objective, valid and reliable instrument. No single test can serve the great number of occupations and the great variations in skill, judgement and information which represent acceptable journeyman competence. While there are several broad categories of tests devoted to specific competency measurement, each has its strengths and weaknesses. The advantages and disadvantages of the various types of tests must be considered when the test form is selected for measuring occupational competence.

I. LIMITING FACTORS IN TEST SELECTION

Occupational competency requires the measurement of many different characteristics, factors or qualifications in different occupations. The choice of the appropriate test is influenced by what the test is to measure and the circumstances and conditions under which the test can be administered.

The three types of performance tests mentioned earlier, as well as several types of written tests, each measures some aspects of competence and has its place in measurement Which type of test to use for a given situation depends upon a number of factors such as those identified under the following five major groupings:

A. The Nature of the Trade or Occupation

1. Does the occupation involve a high degree of manual skill and technical information?

2. Does it require a high degree of technical information and relatively simple manipulative skills?

3. Does it require a high degree of theoretical background and technical training?

B. The Type and Amount of Machinery and Equipment Required

1. Is the machinery or equipment readily available?

2. Is the equipment, machinery and apparatus required usually found in school shops or laboratories, or will industry make it available?

3. Is the equipment, machinery and apparatus up to date (like numerical control equipment, optical measuring equipment, electrical-electronic control devices and measuring instruments in a machine shop)?

4. Are the facilities required readily accessible and available at times convenient for taking examinations?

C. The Materials, Parts and Supplies Needed

1. Is the material required within a reasonable cost range?

2. Can the material, parts and components be readily obtained?

D. The Use to be Made of the Test Results

Is the satisfactory passing of the occupational competency examination to be used:

1. To validate occupational experience?

2. To establish evidence of competence?

3. For advanced credit for further study?

E. The Range of Skills and Information to be Evaluated

1. What special aspects of an occupation must be measured?

2. How are the various levels of competence within one occupation to be established?

3. What are the common elements in a cluster of occupations?

4. What constitutes testing comprehensiveness for a level of competence?

It must be apparent that the development of proper instruments for the evaluation of occupational competence is influenced by many conditions, practical considerations and judgements. The test constructor must remember that the ultimate purpose of the test, or tests, is the evaluation of the present competence and proficiency as a tradesman or technician

II. ADVANTAGES AND DISADVANTAGES OF VARIOUS TESTS

Every measuring instrument contains advantages and some limitations. Advantages and disadvantages associated with manipulative (performance and written) tests follow:

A. Advantages of Work-Sample Tests

1. One major advantage of a performance test is in its apparent realism. Performance tests tend to impress upon the candidate a more practical measure of his ability and achievement and to make him feel that his true proficiency is being evaluated.

2. Performance tests emphasize operational skills. While skills in performing a task can, to a limited degree, be measured indirectly by means of written tests, work-sample tests provide the opportunity for direct measurement.

3. Work-sample tests are relatively non-verbal, compared to written tests. Candidates less facile in writing or reading are not penalized as they may be in written tests. Further, the conditions of testing simulate the conditions surrounding performance on the job from simple directions, drawings, etc.

4. Performance tests can involve a work job as a whole and permit the checking of critical steps.

5. Performance tests are, generally, accepted as a reasonable method of relating testing to job duties and responsibilities.

B. Disadvantages of Work-Sample Tests

1. Work-sample tests are difficult to develop and construct. This is particularly true in highly technical fields; such as, electronics equipment operation, maintenance and repair, technical design and similar fields.

2. Performance tests are not easy to administer. In order to assure maximum value, work-sample tests require time and competent observers, and can be administered to only a few candidates at the same time.

3. Work-sample tests may tie up equipment which may be needed for instructional purposes in schools or production when industrial facilities are used.

4. Accurate and consistent scoring methods are difficult to develop and maintain. Most scoring on performance tests is the product of observation, evaluation and judgement. Inconsistencies and errors in judgement, on the part of the examiner, may influence test results, unless written directions and measuring scales are developed and used by the examiners.

5. Only a limited number of operations or steps can be tested in normal time limits. Preliminary research has indicated that performance on certain selected tests may establish overall competency to a fairly high degree. The validity and reliability of such tests depends upon the careful analysis and selection of those items which are representative of overall competence. Personnel skillful in this area are not readily available.

C. Advantages of Written Tests

1. One of the major advantages of the written test is that it can be administered to a large group of candidates conveniently and economically.

2. Written tests permit a wide sampling of facts, information, knowledge of materials, interpretation of specifications, and other aspects of the occupation.

3. Trade content can be checked accurately through the use of objective-type questions.

4. Written tests permit objective statistical treatment of the test results. Scoring of test question answers can be achieved by clerical personnel working from a preestablished answer sheet.

5. In the preparation of test questions, it is easier to prepare a pool of questions of varying difficulty level for alternate forms of the same test.

6. Written test content can be revised easily.

7. Computer facilities and procedures make it possible to process test results quickly.

8. Written tests can be readily analyzed and researched as to reliability, validity, difficulty level, etc.

D. Disadvantages of Written Tests

1. Occupational skills and procedures involved in the daily work of the occupation cannot readily be measured.

2. Once printed, it is time consuming and costly to make revisions in content or in the type of examination questions.

3. There is major reliance on the measurement of facts and data. Practical judgements and the demonstration of skills, handling of equipment and tools, etc. do not lend themselves to measurement by written tests.

4. Reliance is placed on the individual's ability to read and interpret written material under unusual pressures of time and setting.

5. The planning and preparation of test items requires special knowledge of the occupation to phrase questions in the language of the occupation. Thus, test items are expensive to prepare.

III. OCCUPATIONAL COMPETENCE MEASUREMENT WITH WRITTEN AND PERFORMANCE TESTS

A. Value of Multiple-Choice Test Items for Written Tests

Written tests contain a great variety of different types of test items: true-false, completion, multiple-choice, matching, short answer, and other problem-solving items. Each type of test item has certain advantages and disadvantages.

A great deal of research carried on in recent years indicates that, for the written test, multiple-choice test items give the best results. Well-phrased multiple-choice test items require reasoning, factual information, computational skills, communication skills, judgement, etc. For most tests, this type of question is suggested. Small advantages, other forms may have, are outweighed by the overall results obtained from tests consisting of multiple-choice test items. The many forms of multiple-choice test items and combinations incorporate most of the advantages of other objective types and provide the necessary flexibility to vary the form to meet differing occupational requirements.

B. Need for Both Written and Performance Tests

For effective performance in any occupation an individual acquires manipulative skills and certain "know how" or trade information.

With the rapid development of new machinery, methods and materials, many occupations require an increasing amount of technical information. While skills are ever changing, they do so mostly towards greater intricacy and the application of more information. The cutting tools of the machine shop of today are vastly different from the simple carbon and high speed materials used in earlier years. The tradesman must have the skills of utilizing diamond tools as well as be able to shape and form various cutting angles used for different materials. The electronics service technician must not only be able to test tubes, transistors and solid state circuits, and replace them but, also, he must understand the underlying principles which make them function and trace difficulties to their source when trouble develops.

Evaluating occupational competence requires the development and application of both written and performance tests.

Chapter III

TEST DEVELOPMENT PROCEDURES

I. THE IMPORTANCE OF PROPER PLANNING

The development of tests for measuring occupational competence requires great care, imagination and resourcefulness, as well as sound judgement. It is a time-consuming process. The purpose of occupational competency evaluation is to establish evidence that a tradesman, or industrial technician, has the necessary range of skills and the depth of knowledge required as subject content to be imparted by the teacher in developing other selected students to perform work skills, and obtain technical knowledge essential to employment. Therefore, the first step, in any test construction project, is to define the purpose the test is to serve, set up criteria for the construction of the test, plan the procedures to be followed, and determine the specialists who must be involved. Careful preplanning is essential.

II. OUTLINE OF STEPS IN PLANNING WRITTEN TESTS

A. Sources of Related Job Classifications

Increased specialization in many occupations makes it difficult to recruit the all-round tradesman who entered his occupation via the apprenticeship training program. Special vocational education school programs for youth and adults, inplant training, the pick-up method, and other patterns of training are avenues through which people become proficient in their occupations. To enable people with widely varying backgrounds of preparation to become teachers, a test must sample the skills and knowledge of a cluster of related job classifications. These clusters can be identified from such sources as:

1. The Dictionary of Occupational Titles or other Standard Industrial Classifications,
2. Job descriptions and classifications from employers and occupational studies of labor, management, etc., and
3. United States Office of Education taxonomy of vocational education fields and major occupational groupings in job clusters.

B. Procedures for Determining Occupational Competencies to be Measured

1. Establish the purposes and rationale behind the development of the occupational competency examination.
2. Chart the overall project, as suggested in Table I.
3. Select qualified personnel to serve as a test advisory and development committee.

4. Identify jobs by title and specifications that are typical of the occupational area and give the range of competencies required for the whole occupational area. For instance, Table II identifies selected job titles for the electronics industries occupations (communications) that range from single skill operatives to highly skilled professional persons.

5. Group the job titles, which require similar skills, technical knowledge, and other occupational responsibilities and duties, into job clusters on different levels (Table III).

6. Determine the levels (according to objectives to be achieved by the test) for which the occupational competency test is to be developed.

7. Make job analyses of jobs in the cluster in term of:

   a. Manipulative skills

   b. Trade theory and judgements

   c. Trade communication (technical language specifications, drawings, etc.)

   d. Computational skills (mathematical calculations and reference tables)

   e. Applied principles of sciences

   f. Technological knowledge needed

   g. Industrial safety, health and hygiene practices

   h. Other information essential to the occupation

C. Job Description and Test Specifications

1. Prepare job descriptions. These should define job skills and accompanying trade theory for the performance part of the test. Related trade/technical information, knowledge, judgements, and other occupational competencies the potential teacher must know should be described as a guide for the written part.

   Table IV describes special occupational items for which an occupational competency examination is to be prepared for the electronic industries occupations (communications).

2. Establish the scope of the test in terms of level and range of skills and information to be evaluated.

3. Determine the nature and format of test items that should be prepared for the written part of the test.

4. Set up a specification table to roughly establish the component parts of the occupational competency test for each main and subdivision. The number of proposed test items to consider should be noted.

   The structure of Table V suggests columnar groupings for which test items should be prepared.

D. Alternatives for Preparing a Job Description to Establish Test Content

At times, it may not be possible to carry out a job cluster analysis and complete the required amount of detailed search and analysis. Even under restrictive conditions, a job description should be prepared. While it may not be as comprehensive, the description will help to identify test items that are not easily determined by other means.

The following techniques may be used to prepare a job description:

1. Actual on-the-job on-site observations may be made of the work to be done by a trained person in the occupation. Such observations are made by occupationally competent persons, together with a test specialist.

2. A representative group of experienced craftsmen within the occupation may prepare a job description.

3. The job description of the job cluster may be prepared as described earlier.

In any case, the job description is another tool which provides direction and assistance to the test developer.

E. Test Items, Test Construction and Evaluation

1. Establish procedures for electronic data processing relating to test construction, administration, scoring, recording, analysis and evaluation.

2. Assess the test structure and the number of test items needed to achieve the test objective.

3. Prepare a pool of test items with answers.

4. Evaluate each test item and revise as needed.

5. Arrange test items in proper sequence within the occupational competency test.

6. Prepare a Handbook for examiners with answer sheets.

Table I SEQUENCE IN TEST CONSTRUCTION PLANNING AND PERSONNEL REQUIRED
ELECTRONIC INDUSTRIES OCCUPATIONS-COMMUNICATIONS

Levels of Job Clusters	Analysis of Job Titles in a Cluster	Test Specifications Derived from Objectives and Analysis	Test Construction and Development
Single Skill A	Major Divisions Subdivisions	Test Purpose	Format of Examination
Semi-Skilled B	Job Content Analysis	Test Standards	Arrangements of Test Items Number of Test Items Alternate Forms Method of Scoring Rating, Evaluation and Reporting
Skilled Trade C	Manual Skills Trade Practices Trade Information Related Mathematics	Test Utilization	
Skilled Technician D	Related Science Technical Language Drawing and Representation	Test Scope Job Clusters Job Description Skills Job Knowledge Related Information	Methods of Testing
Semi-Professional E	Trade Judgement Trade Safety Industrial Hygiene Industrial and Labor Relations	Test Type Test Evaluation Scoring Ratings	Population Sample Form of Validation Reliability Comprehensiveness Test Item Pool Types and Number of Test Items Types and Number of Performance Jobs
Professional F			

Selected Personnel Involved in Test Development

Teacher Trainer Industrial Specialist Manpower Analyst Trade Teacher	Trade Teacher Manpower Analyst Teacher Trainer	Trade Teacher Industrial Specialist Test and Measurement Specialist Teacher Trainer	Trade Teacher Industrial Specialist Test and Measurement Specialist Teacher Trainer Other Psychological Testing Service

Table II. SELECTED JOB TITLES FOR ELECTRONIC INDUSTRIES
OCCUPATIONS - COMMUNICATIONS*

Job Titles	D.O.T. Code
Electronics Assembler	726.781
Electronics Assembler Development	726.781
Electronics Mechanic	828.281
Electronics Technician	003.181
Radio Mechanic (and ind.)	823.281
Radio Mechanic - Troubleshooter	720.281
Radio Mechanic - Repairman	720.281
Television - Installation	823.781
Television - Service & Repairman	720.281
Television - Chassis Inspector	720.687
Television - Transmitter Assembler	726.884
Electronics Sensing Equipment Assembler	716.884
Component Inspection Technician	828.281
Customer Engineering Specialist	828.281
Radioactivity Instrument Technician	828.281
Electronic Mechanic Apprentice	828.281
Electronic Sound Technician	024.288
Public Address Serviceman	823.281
Electronic Technician	003.181
Systems - Testing - Laboratory Technician	003.181
Radio Engineer)	003.081
Television Engineer)	
Radio Equipment Installer	823.281
Radio Repairman Domestic	720.281
Radio Mechanic Helper	823.884
Electrical and Radio Work-up Man	825.381
Television Repairman Apprentice	720.281

T.V. Audio Man)	F.C.C. 1st class license required
T.V. Transmitter Maintenance)	
Microwave Technician)	F.C.C. 2nd class license required
Mobile Communications Equipment)	
Two-Way Radio Communications)	

* Derived from the Dictionary of Occupational Titles

Table III. JOB CLUSTERS FOR FIVE MAJOR LEVELS OF OCCUPATIONAL PROFICIENCY FOR ELECTRONIC INDUSTRIES OCCUPATIONS (COMMUNICATIONS)

Level I	Level II	Level III	Level IV	Level V
Electronic Assembler 726.781	T.V. Installation 823.781	T.V. Service and Repair 720.281	Transmitter Maintenance F.C.C. 1st Class License	Radio Engineer T.V. Engineer 003.081
Electronic Assembler Development 726.781	T.V. Chassis Inspector 720.687	Component Inspector Technician 828.281	Mobile Communications Equipment F.C.C. 2nd Class License	Electronics Technician 003.181
Electronic Mechanic Apprentice 828.281	Electronic Sensing Equipment Assembler 716.884	Radioactive Instrumentation Technical 828 281	Microwave Technician	Systems Testing Laboratory Technician 003.181
Radio Mechanic 823.281	Radio Equipment Installer 823.281	Electronic Sound Technician 024.288	Two-Way Radio Communications Technician	
Radio Mechanic Helper 823.884		Public Address Serviceman 823.281	T.V. Transmitter Maintenance	
T.V. Repairman Apprentice 720.281		Radio Repairman Domestic 720.281	Transmitter Assembler 726.884	
		Electronic and Radio Work-Up Man 825.381	Customer Engineering Specialist 828.281	
		Radio Mechanic Troubleshooter 720.281	Sound Technician 024.288	
		Radio Mechanic 823.281		
		Electronics Mechanic 828.281		

Table IV. JOB DESCRIPTION (OCCUPATIONAL COMPETENCIES) FOR ELECTRONICS OCCUPATIONS COMMUNICATIONS

General Comments

For an electronics industry occupational competency testing program to be used nationwide, it is essential to include test items of common use in various sections of the country rather than specialized practices of a particular area. The competent tradesman should be able to perform his work regardless of geographic location or national origin. The following job description should be used as a guide for test development:

Job Duties: Skills and Related Trade Theory

- Repairs, adjusts and maintains electronic equipment such as: radio receivers (both AM and FM), audio amplifiers, tape recorder units; communication receivers, transmitters, receivers; T. V. receivers.

- Must locate faults, replace defective components, test, align and calibrate equipment for: power output, correct frequency, distortion levels, audio and video quality; measure current, voltage, resistance wattage and frequency; test all components, test tubes, diodes and transistors; trace circuits and make all necessary corrections. Installs, tests, adjusts and modifies inter-communications systems, using the instruments of the trade.

Technological Information and Knowledge and Trade Judgment

- Must know electrical fundamentals of AC/DC capacitance, resistance and inductance, series circuits, parallel circuits, series-parallel circuits, resonance, magnetic theory, transformers.

- Must know electronic fundamentals of solid state diodes, transistors, vacuum tubes; diode, transistor and vacuum tube characteristics curves, power supply circuits, A.F. amplifier circuits, RF amplifier circuits, video amplifier circuits, microwave circuit techniques, oscillator circuits, amplitude modulation, frequency modulation, pulse modulation, negative feedback circuits.

- Must know troubleshooting methods; signal injection; signal tracing, voltage analysis and resistance analysis techniques.

- Must know basic power distribution system; 117-volt single phase, 220-volt single phase; two and three phase systems.

- Trade judgment. Must exercise judgment in the selection of the most suitable procedures for the location of faults, in the selection of the appropriate test equipment, in the repair techniques and the replacement of substitution of parts of modular units when needed.

Table V. TEST ITEM PLANNING FOR INDUSTRIAL ELECTRONICS-COMMUNICATION OCCUPATIONAL COMPETENCY EXAMINATION

(Suggested Columnar Headings Under Which Skills and Related Technical Knowledge are Identified. - Not to Scale)

Main Division	Subdivision	Trade Skills	Trade Theory	Technological Information	Trade Judgment	Trade Computation	Trade Science	Trade Communication	Industrial Safety and Hygiene	Other Related Information	Total of Test Items

7. Review the test in committee for structure, time, relevance, etc., utilizing the services of a test and measurement specialist.

8. Determine a representative sample group for pilot testing the examination.

9. Field test, evaluate and refine.

While not complete and detailed, the basic steps outlined in this Section II must be followed. Only after the test items have been field tested can final revisions be made. Weak, inadequate or duplicate test items can be eliminated. Thereafter, an instrument can be developed that gives reasonable assurance of being valid and reliable.

III. PLANNING PERFORMANCE TESTS

A. General Considerations

The steps involved in planning the performance test are somewhat similar to those that were outlined earlier. However, since work sample performance tests are usually concerned with some overall job situation, its scope, content and requirements warrant separate consideration.

Some individuals believe that any performance in which tools are used, a machine is operated, parts are checked and replaced, or things are manipulated, represents a satisfactory performance test. Such a test may seem to have a great deal of face validity, since it appears to duplicate the job situation so well that there seems little doubt that it measures the intended purpose.

This may not really be so. Unless great care is taken and use is made of the job cluster analysis for those skills and methods which are vital to the job performance, and provision is made for step-by-step procedural planning, standardization of administration, and objectivity in scoring, the test can easily become an unreliable instrument. As such, it reflects merely the judgment of the administrator and it may not be valid.

B. Steps in Planning a Performance Test

The following steps will lead to the design of an effective test regardless of whether the performance test involves an actual work job, a simulation situation, a combination of these two methods, or recognition procedures:

1. A work sample must be made up of representative elements of the job titles in the job cluster.

2. Each element in the work sample should permit variation in performance and be applicable to other work jobs.

3. A variety of work samples should be tried. Statistical analysis should be used to determine which elements discriminate, and that no important phase of the job has been eliminated in the process of condensing the test.

4. The tasks selected for the work sample must be representative of the job cluster and job level and require sufficient time to lead to reliable scores.

5. The tasks chosen should command respect and establish the confidence of the individual engaged in the job as being representative of the occupation.

6. The materials, tools and equipment should be reduced to the smallest practical quantity and should be capable of standardization so that the same test may be given under uniform conditions.

7. The performance should involve as little repetition of identical procedures as possible.

8. A preliminary tryout of performance tests should always be made using experts as students to detect possible problems or difficulties.

IV. EVALUATING THE PERFORMANCE TEST

A. Work Method or Finished Product

In evaluating a performance test, the question often arises as to which is more important; the finished product or the work method used and the work habits applied to the completion of a job. For most occupations, the either/or approach is not applicable. Each occupation has its own characteristics. There are great variations in skill and knowledge among the many occupations for which potential teachers must be recruited. Careful analysis is required as to which approach is more important and should receive greater weight.

For example, if the occupation involves troubleshooting of electrical circuits, it is possible to accidentally stumble upon the trouble and then correct the defect without the systematic process of testing and elimination. The equipment may perform according to specifications without the examinee really having demonstrated that he possesses the skills of analysis and step-by-step checking. On the other hand, in the production occupations (machine trades, sheet metal trades, etc.) the nature of the work and the finished product make "accidental aids" virtually impossible. It is easier to evaluate the finished product objectively than to appraise work method and work habits.

B. Identifying Craftsmen Characteristics

Although occupational competence tests are not intended to evaluate and predict teaching potential, it is important for the tradesman and future teacher to demonstrate those work habits and procedures which distinguish the good journeyman from a less competent individual. Eventually, these are the very things he will have to teach and they, thus, represent an important area of competence for his future career.

Therefore, the work sample selected for the performance test must permit an evaluation of work habits as well as permit an objective evaluation of the finished product. The choice of the work sample is further influenced by the length of time required for its completion, the amount of equipment required and accessible, availability of testing personnel and testing facilities, funds, etc.

Equally important is the choice between a single long job which involves selected skills and methods of procedure or a number of smaller assignments with each requiring special, vital skills. Again, the nature of the occupation will largely determine this choice. It will be easier to design a work sample which does not need to be fully completed for the machine trade than for a service occupation. For example, the skill of using layout procedures, operating a drill press, a grinder and a miller can well be combined in one work sample. For another, it is possible to combine lathe work, heat treating and grinding. A single work sample arrangement would require more bulky equipment of various kinds and be thus more limited.

In any case, both method and product must be evaluated. The nature of the occupation determines where the greater emphasis should be.

V. PARTS OF THE OCCUPATION REQUIRING EVALUATION

A. Aspects that Represent Typical Jobs

The parts of the job to be measured in a performance test are those which represent the occupation as a whole. They should be selected on the basis of how critical they are to performance in actual wage-earning employment for the proper operation of machines, apparatus, equipment, etc. Many tasks in a job may be of a routine nature and a great deal of time may be devoted to them in the shop and laboratory.

In the interest of time and economy of test administration, it is important to reduce or eliminate repetition or duplication of certain operations and concentrate on those more critical on the job. For example, after a lathe job has been completed and the examinee is required to demonstrate his skill on the milling machine by cutting a gear, it would be a waste of time and useless repetition to have him complete the whole gear. Cutting three or four teeth would demonstrate whether he knows how to do it and will enable the examiner to measure the quality of the work.

Similarly, in a test of radar maintenance, including the calibration of range marks for search equipment is unnecessary because search type of radar can be operated under combat conditions without close-range calibra-

tion. The test might, also, omit determination of band width by the use of sweep generators because band width of most radar sets is usually fixed and cannot be changed without changing components.

On the other hand, power measurements must be included because the maximum range of the radar set is partially determined by the amount of power transmitted. Unless a repairman can determine the amount of power transmitted, he cannot know whether the weak or few video signals received are due to a defect in the transmitter. In like manner, the test should include the use of a signal generator because the radar repairman needs to know if the radar set is able to pick up and display weak signals from the maximum range of the set. Both of these tasks determine the performance figure of the radar set and are crucial to its return to the proper operation in the shortest possible time. [12]

B. Establishing Critical Items to be Measured

The items to be measured must be derived from the occupational analysis. This should facilitate the determination of the operations to be included in the work sample. The analysis chart would show the repetition of various operations in the different job clusters. It becomes then a matter of judgement to distinguish the critical and important items from those that are common to the various job classifications and less important for inclusion in the test.

VI. TERMS DESCRIBING AREAS FOR SPECIFIC TEST CONTENT

A. Variations in Test Terms

Repeated reference has been made to job analysis, and several approaches have been described. Some terms, found in the literature, are confusing. One author, for instance, calls the breakdown of the steps in thread cutting an analysis of a job; others refer to the same item as task analysis. Still others use such terms as cognitive and perceptual objectives for making an analysis. [13]

Under the cognitive approach are such subdivisions as knowledge, understanding, application of knowledge, application of understanding. The perceptual approach covers such areas as sensation, figure perception, symbol perception, and perception of performance. [14] A further distinction is made by Simpson, who discusses the psychomotor domain. [15] Moore approaches the problem in terms of a proposed taxonomy of the perceptual domain. [16] Krathwohl discusses the affective domain. [17] While these approaches are different in terms and essential toward developing new approaches to teaching and the evaluation of learning, they are, at this stage, predominantly academic approaches. For aiding the test instructor, faced with the immediate task of developing of occupational competency examinations simpler definitions are needed.

B. Terms Defined

The following simplified definitions of terms are used in this Handbook. Examples of each of the four major terms are given to identify applications in a number of different occupational settings. Content relating to each term is derived from the job analysis.

1. Operation - a basic skill involving the correct manipulation of tools, machines, equipment, apparatus and supplies to complete a specific task according to procedures and practices followed in a trade or industrial/technical occupation.

   a. Laying out object lines to describe a part or mechanism

   b. Straight turning a work piece on a wood lathe

   c. Testing a series circuit for voltage

   D. Assembling rings on a piston

2. Related Technical Information - information necessary to carry on the work of the trade, involving basic principles of mathematics, science, drawing, etc., applied to practical trade practices.

   a. Electronic theory, as applied to transistors and solid state circuits

   b. Principle of combustion, as applied to power systems

   c. Mechanical advantage, as applied to pneumatic and hydraulic clamping devices

   d. Theory of atmosphere control in heat treatment of metals

3. Trade Theory - involves technical content knowledge, directly related to an operation.

   a. Letter and number drill size systems

   b. Grinding wheel specifications and surface finishes

   c. Transistor functions, capacities, structure

   d. Determining parts from m/gn. catalogues for a certain year car model

4. General Trade Knowledge - this term refers to ancillary content, work attitudes and personal relationships.

   a. Machine and tool safety

   b. Content bearing on health and industrial hygiene

   c. Employer, employee, and community relations

VII. SUMMARY EXAMPLE: CONSIDERATIONS FOR THE DEVELOPMENT OF WRITTEN AND PERFORMANCE TESTS FOR ELECTRONICS INDUSTRIES OCCUPATIONS-COMMUNICATIONS

A. Test Purposes: To provide measures of the technological knowledge and ability and the trade skills through which to establish the occupational competency of an individual. The test results may be used alone or in combination to:

1. admit examinees to industrial teacher education programs,
2. to determine the extent to which college credit should be granted when pursuing an undergraduate or graduate degree program,
3. diagnose occupational strengths and weaknesses, and
4. certify capable craftsmen with professional preparation as vocational teachers.

B. Test Standards: General -

The tests must represent the best possible instruments acceptable for validity, reliability, comprehensiveness, currency and economy of administration. The final instrument should evaluate occupational competence for several levels of job clusters--each on the achievement level requirements of industry.

The job clusters, including the semi-professional technician level, represent the range and levels for which occupational competency examinations should measure the potential teacher's competency.

C. Methodology: General -

The final tests should be based on an occupational anaysis. Job titles in occupational constellations in the electronics industries-communication and related occupations should be clustered according to various levels of competency. The levels are established from operative to semi-skilled craftsmen, skilled craftsmen, technician in industry service, to semi-professional and professional engineers.

D. Approach to Occupational Competency Test Development:

1. The job clusters at each level require an analysis of common elements. These elements involve manipulative skills, technology, mathematics, principles of science, communication (drawing, reading of specifications, etc.), design, and materials involved in the job performance of the occupation.

2. The analysis must include not only trade skills and knowledge but, also, knowledge of equipment, materials, apparatus and tools. These, when applied to the work of the occupation, need to be evaluated with a job performance test.

3. The informational skills, trade knowledge and areas of work involving judgement can be evaluated through written examinations.

4. The major part of each examination - written and performance - should adequately represent significant learning areas of the various job clusters. Collectively, the examination should be comprehensive in scope. The test items, in each part, should be mostly objective and in sufficient number to assure the broadest possible measurement of occupational competency.

E. Test Utilization: The achievement levels on these tests (for admission to Industrial Training Programs, and/or for preliminary state certification, and/or advance credit for a program of study for a degree) will be a discretionary matter for the various states and industrial teacher training institutions to determine.

F. Test Scope - Electronics Industries Occupations-Communications:

Proficiency in the skills, information, methods, procedures, and judgements of the Electronics Occupation-Communication are a fundamental requirement for an individual interested in entering the field of teaching in vocational education. The proficiency examination must be designed to:

1. Evaluate the applicant's skill in operating the various instruments: such as, signal generators, oscilloscopes, meters, and of handling the necessary tools and materials commonly used by the competent tradesman in the field of electronics.

2. Evaluate the applicant's knowledge and information necessary to formulate judgements involved in the planning and carrying out the work of the occupation.

3. Evaluate the applicant's skill and knowledge in applying the methods and procedures expected of a competent tradesman in the field.

4. Establish sufficient and professionally acceptable evidence for granting of undergraduate credit towards a degree in Teacher Training Institutions or for advanced certification.

G. Job Description: Electronics Industries Occupation - Communications:

1. General Comments: The electronics industries include a number of jobs differing in detail and difficulty level which, nevertheless, contain substantial common elements. Some are narrowly specialized; others require a broad background of information and skill. The prospective teacher must have the occupational competence to teach preemployment classes involving fundamental skills and concepts; advanced modern trade knowledge, skills and theory; preemployment classes for special industrial employment; occupational extension classes to extend the work skills and technical knowledge of the skilled worker; apprentice classes; and foremen and supervisor classes.

   The Occupational Competency Examination must measure the individual's occupational competency required to teach at all levels within secondary schools, post-secondary institutions, and adult education programs.

2. Earlier comments and illustrations (refer to the preceding Tables I through IV) related to job cluster practices on multiple levels of employment and other job characteristics and descriptions. The job description, naturally, includes job duties, skills and related trade theory, technological information and knowledge, and trade judgement.

H. Occupational Competency Measuring Instruments:

1. Written Test : Knowledge of trade information, properties of materials, principles of design, and operating principles of equipment; computational skills; and the abilities to interpret blueprints, wiring diagrams, schematics, and specifications, to apply scientific principles, and to exercise trade judgement can be evaluated through written tests.

   It is important that the questions be objective in nature with correct answers provided. The number of test items should be sufficient for each job cluster to adequately measure the individual's competence.

2. Performance Test: The varying nature of service, maintenance and repair make the choice of jobs for evaluating a craftsman's competence difficult. Every job to be done should involve essential key skills without which a job could not be completed. In the electronics field the performance test may involve a selection of short jobs on different kinds of apparatus, or a single job with a number

of difficulties involving the use of meters, equipment, tools, etc. Thought should be given to practicality, facilities, and the availability of units on which work can be done. At all times, the jobs selected for performance and evaluation should involve a range of practical skills, which, when properly performed, establish the occupational competence of the prospective teacher.

I. Scoring Methods:

1. Written Test: The scoring of written tests should be accomplished by clerical persons. A special key form that permits the identification of right or wrong answers will be provided. It is preferred to develop a test that can be scored by electronic data processing equipment. Scoring must be done efficiently, accurately and confidentially.

2. Performance Test: Evaluation of the tradesman's occupational competence involves consideration of:

   a. The skill with which he handles the instruments, apparatus, and hand tools of the technical occupation.

   b. The evaluation should take account of the working methods and procedures of a candidate and the performance of the completed unit.

   c. The skill he demonstrates in using reference materials, technical manuals, specifications and directions.

3. The evaluation should be made on troubleshooting, checking and tracing circuits, tracing various stages of a unit, identifying the trouble, replacing a defective part or unit, aligning, calibrating or testing a unit for proper performance.

4. Items for Skill Evaluation of the Examinees:

   a. Preparation and set-up of units to be serviced

   b. Arrangement of instruments, meters, tools to be used

   c. Orderly and logical procedures from unit to unit

   d. Steady work - step after step....occupational confidence

   e. Personal and machine tool safety

f. Neatness in working

g. Recording of all necessary information

h. Checking measurements carefully and accurately

i. Demonstrating capacity to use reference materials properly

j. Checking proper performance of the test job

J. Rating Form: A rating form should be designed which permits unobstrusive recording of specific observations; contains numerical or letter values for the items to be observed, and permits easy and quick computation of test results.

Chapter IV

CONSTRUCTION OF TESTS

I. PERSONS INVOLVED IN TEST CONSTRUCTION

A. The Role of the Subject Specialist

Unfortunately, there is no universal agreement as to what constitutes vocational success, which is the goal of prediction in the case of Civil Service tests. What is accepted as evidence of success is largely a matter of the combined judgement of a number of competent persons as to what variables to include in establishing adequate measures.

The test developer knows the principles of personnel measurement; appropriate methods and techniques for determining whether a test serves the purpose for which it is designed, and whether test results provide accurate and reliable evidence of occupational competence. His knowledge and understanding of the occupation may be of the "talking knowledge" kind.

It is necessary, therefore, to engage the assistance of an individual who "not only knows" his trade but has mastery of the essential skills. In short, he must be occupationally competent. He must know which skills and knowledge are vital to the performance and the main problems encountered in his occupation and at what level.

The subject specialist provides the occupational content for the test. He makes the job analysis, plays an important role in the content of the test items, and assists in establishing what constitutes a measure of occupational competency.

B. The Role of the Vocational Teacher

Recruiting tradesmen experienced in test development is not an easy task. Availability, time and cost are often limiting factors. Vocational teachers who are presumed to be masters of their occupation with an understanding of test practices can provide valuable assistance. Under proper guidance and direction, they can make a valuable contribution to the test content.

However, a word of caution is in order. Since teachers develop tests for use with their own students, they tend to emphasize instructional content rather than prevailing occupational practice. Concerned with teaching methods and procedures, they frequently recommend content which they believe a tradesman should have rather than what may actually be required on the job. Often, they include subject content they, themselves, have acquired after they started teaching. Furthermore, there is a tendency towards developing test items that may be somewhat "schoolish" in language and application.

C. The Role of a Committee of Specialists

1. Function of the Committee

The process of developing and constructing occupational competency tests is a cooperative undertaking. Representatives from the occupation or subject specialists need to be brought together. In selecting such individuals, it is important to have representation from various branches of industry in which the trade is practiced. This reduces the possibility of the resulting examination being biased in one direction.

For example, in electronics communication, it would be unwise to have only representation from radio and television service maintenance. Similarly, in the machine trade, it would be inviting bias if only tool and diemakers were represented and no machinists included. It is of equal importance to have representatives from several different manufacturers rather than one large organization. Such a committee should include teachers of the occupation for which a test is being developed.

The committee's major function is to determine the subject content, areas to be covered and the development of each item. The test specialist serves to recommend the test form and the technical test merits of each item. Since occupational competency deals with specialized content and skills, each test item should be prepared by the subject specialist. This first draft should then be appraised for technical defects by the test specialist.

An item which covers an important skill, fact, or judgement may still be unsuited for differentiating between a competent, or less competent individual if it is phrased ambiguously or it is too familiar or unfamiliar to the group being tested. Close cooperation between subject and test specialists is important in every step of test development. [18]

2. Size of Committee

The size of the committee is critical. If it is too small, there will be inadequate representation of all facets of the occupation. Too large a group may slow down the process, increase developmental costs and be cumbersome. While it is possible for one subject specialist and one test specialist to construct a test, it is more advantageous to pursue the following in organizing a committee:

a. The subject specialists should represent several areas of the occupation.

b. Vocational teachers should be represented.

c. The composition of the committee depends upon the nature of the occupation and difficulty, or complexity, of its content. There should be a balance among the representatives.

d. Working groups tend to become less efficient and productive as the size increases.

e. While it is not possible to give definite numbers, experience has shown that groups of 3 to 6 function well in developing an occupational competency test for a major occupational area.

f. Should a larger group be needed, their output becomes more effective if the group is subdivided into smaller working groups.

3. Committee, Direction and Working Procedures

A committee makes valuable contributions toward test development, as the broader experiences represented should assure wider coverage of a given field. However, there is always the hazard that in a large group much effort is spent in fruitless discussion. It is important that some preliminary work be done by the test specialist.

The great variety of occupations make it nearly impossible to establish a rigid procedure. Care must be taken that any preliminary plans do not limit the efforts of a committee. At the same time, certain directions assure more effective productivity.

The actual work expected of the committee, which may include the following, should be established beforehand to assure greater productivity of the committee's efforts.

a. The job analysis must be reviewed.

b. Test specifications are to be reviewed, revised or actually established.

c. A test grid, or pattern, or outline must be established as a base for deriving the number of test items needed for each subdivision of the occupational job cluster.

d. Procedures for developing test items must be established.

e. Evaluation must be made of factors established for judging test results; written and performance.

f. Suggested contacts in industry for testing the examinations are to be identified.

g. Recommended standards to be employed for judging test achievement are to be developed.

II. FUNCTION OF TEST SPECIFICATION

A. Range of Test Specifications

A test specification may be a comprehensive covering of test purpose, standards, methodology and test utilization. It may include the scope, job description, job cluster, types of tests desired, scoring methods and items recommended for evaluation. It may also include a topical outline

recommended for evaluation. It may also include a topical outline listing the trade content for written and performance evaluation and an estimate of the number of test items recommended for evaluation.

Table VI.--SPECIFICATION FOR WRITTEN TEST - ELECTRICITY (Partial)

Main Divisions	Factual Knowledge	Application of Knowledge	Understanding of Processes	Calculations	Technical Interpretations	Total Test Items
D.C. Circuits						
Series Characteristics	2	3	2	3	8	18
Parallel	2	3	2	4	8	19
Combinations	2	6	5	8	8	29

Other forms of specifications may substitute processes, or operations, or skills. Table VI merely indicates a pattern from which can be derived the estimated number and application of test items that are to be developed.

Many other variations of test specifications are possible. In academic tests, such a specification is often called a Test Plan in which objectives and different subject areas are listed (Table VII).

Table VII.[19] TEST PLAN LISTING OBJECTIVES AND SUBJECT AREAS

Objectives	Subject Areas		
	Foods	Digestion	Assimilation
Knowledge of Terminology			
Knowledge of Facts			
Knowledge of Trends and Sequences			

Where it is desirable to establish different levels of occupational competence, a test specification (also called an outline) may be helpful (Table VIII). The grid in a multiple level outline shows some common and some unique areas that must be tested.

Table VIII. MULTIPLE-LEVEL OUTLINE

Areas of Knowledge and skills	Novice	Apprentice	Journeyman	Master
I	25	16	10	--
II	10	10	10	10
III	10	15	15	15
IV	20	15	10	10
V	20	20	20	20
X		10	15	15
XI			10	15
XII			5	20
XIII				10
XIV				20
Total Test Items	85	86	95	135

Whichever form is employed for planning an examination, it must be apparent by now that the comprehensive specification provides the overall pattern. Eventually, it will be necessary to develop a more specific blueprint to serve as a reference and pattern for the number of test items required for each category. A test developed in accordance with definite specifications is more likely to yield a score that is meaningful than that of a test whose items are assembled at random.

III. FORMULATION OF TEST ITEMS

A. Choice of Test Items

Are essay questions more reliable than objective-type test items? Are multiple-choice items better than true-false items? The relative merits of different types of test questions, or items, were long a popular debate in educational circles. We now know that there is no general answer to such questions. There is no one type that is inherently superior to other types. How well an item is constructed is more important than the type the item represents. Each type of item, moreover, has certain strengths and weaknesses. The choice of item types for classroom examina-

tions, therefore, evolves from the special advantages and disadvantages for each of the purposes the teacher has in mind."

B. Popularity of Multiple-Choice Test Items

For purposes of test construction for occupational competence, factors of economy of administration, effectiveness of scoring and statistical treatment of data are important considerations. Among the many objective-type tests, the multiple-choice has gained more and more popularity. The multiple-choice form is equally as good as the true-false or short answer item for measuring factual knowledge. Furthermore, it has been found to be excellent for measuring understanding and for the ability to apply concepts and knowledge to unique situations.

The military and civilian services are largely committed to the use of multiple-choice items, since they consider multiple-choice items superior to other types. Since the multiple-choice type is adaptable to the best answer approach, it avoids the ambiguity associated with the application of a standard of absolute truth, which is the chief weakness of true-false test items. Also, multiple-choice type items are independent of each other, so the problem of finding a number of parallel relationships, which frequently cause difficulty in the matching type of items, is avoided.

Other advantages follow. If adequate checks of accuracy of statement are made before the items are used, the multiple-choice form will usually avoid unforeseen difficulties due to the subject's misunderstanding of the intent of the item.

Multiple-choice items permit the use of standard machine scorable answer sheets. Since most testing programs are equipped to deal with multiple-choice answer sheets, standardization in a number of different establishments can be undertaken as routine practice.

C. Changing Alternative Types of Test Items

Newer test item types include changing-alternative items. These require a best answer to be selected from several alternatives which change with each item. The most usual form is the multiple-choice item.

The changing-alternative type of item is adaptable to testing higher mental processes, such as inferential reasoning, fine discrimination, and knowledge of isolated facts. It is the most flexible of its kind of test item available for varying types of mental processes on the basis of subject matter.

D. Summary of Objective-Type Tests

For all its advantages, there are, nevertheless, occasions when other types of objective test items may serve a definite purpose. Table IX summarizes the more commonly utilized objective test item forms, their advantages and disadvantages, and where they may be used in occupational competency testing. (Table IX, pages 47-49)

Table IX. SUMMARY OF COMMON OBJECTIVE-TYPE TEST ITEM FORMS AND CHARACTERISTICS

Name of Test Item Form	Advantages	Disadvantages	Possible Use
• Selection • True-False • Constant-Alternative • Right-Wrong	• Large number of items can be administered in relatively short time • Wide sampling possible • Entirely objective	• Difficult to create statements that are universally true • Susceptible to guessing	• For situations when the choice is either-or • Things one must know to make a single choice • Testing facts rather than reasoning • Immediate recall
• Selection • Matching Items	• Compact, requires relatively little reading time • Small guessing factor	• Not as flexible as other test items • Restricted to situations in which material lends itself to the listing of a number of important and related concepts • Many items cannot be conveniently grouped • Requires greater care	• Rapid survey of specific aspect of subject matter • Sequence of steps in a process or operation • Recognition or identification of parts and relationships

Name of Test Item Form	Advantages	Disadvantages	Possible Use
• Supply Type • Short Answer	• Eliminates guessing • A natural form of question • Easily prepared	• Scoring not completely objective • Type of item relatively unfit for testing	• Suited for problems • Mathematical and physical science concepts, sketches, diagram reading
• Supply-Completion	• Eliminates guessing • Easily prepared • Useful in problem solving situations (mathematics and physical science) where results can be expressed in a few symbols • Verbal facility and memorization of facts	• Tend to be ambiguous • Scoring not as objective as selection-type test • Expert judgment required to evaluate some answers when different from scoring key • Tends to be confined to identifying, naming and associating facts. Relatively unfit to test understanding	• Testing terminology • Recognizing and recalling information and specific facts • Oral trade test items: names, parts, etc.

Name of Test Item Form	Advantages	Disadvantages	Possible Use
• Picture Trade Test	• Admits more detailed test items • Problems simply stated through reference to a picture • Differentiates better between apprentice, journeyman and master • Closely related to trade • Tradesman has more confidence in test o Permits more intricate test items o Less subject to coaching o Picture often calls forth information not otherwise possible	• Cost of production greater • Awkward to give • Danger of unfamiliarity with item used by tradesman • Pictures require continuous updating	• Identification of parts, objects • Construction of component parts • Relationships of components and parts • Tools, instruments, materials used o Circuits, Diagrams, Schematics
• Multiple-Choice	• Most useful and flexible of all test items o Measures understanding, concepts and knowledge; inferential reasoning • Guessing reduced with increasing number of choices o Scoring objective	o Difficult to construct o Time consuming to develop o Does not permit organizing a solution as in a written test solution	o Nearly all phases of the occupation except where organization of descriptive or explanatory problems are involved

IV. VARIATIONS AND USE OF MULTIPLE-CHOICE TEST ITEMS

For this Handbook, the multiple-choice test item is recognized as the most suitable form. Thus, this portion will be devoted to a discussion of the various ways in which multiple-choice items can be used, how they are best formulated and the various difficulty levels established.

The multiple-choice item consists of: (1) an introductory question or incomplete statement called the stem, and (2) two or more responses from which the one best, or correct, answer must be selected. Many variations in approaches possible can be seen from the following examples and descriptions:

A. The Stem as a Question or Incomplete Statement

The stem of multiple-choice items may be classified according to the form of a question or an incomplete statement. This is illustrated in the following two examples:

1. Question-form example:

A square has how many sides? (Stem)

a. Five
b. Six
c. Three)
d. Four) Responses

2. Incomplete-statement example:

The number of sides contained in a square is (Stem)

a. Five
b. Six
c. Three)
d. Four) Responses

B. Responses Requiring Selection of One Right Answer or the Best Answer

Multiple-choice items may be classified as to the "one right answer" or the "best answer" type. When an incomplete statement is used care must be taken to ensure that it is equivalent to a direct question.

1. The one-right-answer type

In the one-right-answer type, the trainee will be required to identify the one correct response listed among a number of others that are totally wrong. In writing this type, special care must be taken that there is only one correct response in the alternatives. Example:

An auger bit that bores a 1/2-inch hole is stamped with what number?

a. 6
b. 2
c. 8
d. 4

2. The best-answer type

This variation will require the trainees to select the best response from a series of alternatives. The use of this form of item helps to measure judgement, or inferential reasoning, or complete understanding required in selecting the best response. This type of item may be varied to require the trainee to indicate the worst response or the least desirable solution.

Example: You are finishing a fine piece of furniture and notice a bad spot in the wood. Which of the following methods would be the best to use in repairing the surface?

a. Fill the hole with stick shellac
b. Inlay with the same kind of wood
c. Fill with glue and sawdust from the same kind of wood
d. Drill out the bad spot

C. Types of Multiple-Choice Items

The following information classifies types of multiple-choice items. It is hoped that this classification, along with the examples, will give valuable suggestions for developing ideas into test items:

1. Test items relating to definitions and terms

a. What term means the same as . . . . . ?

b. Which of the following statements expresses this concept in different terms?

Example: Erythrocytes are also known as

a. polymorphonuclear neutrophiles
b. square, nucleated cells
c. red blood cells
d. lymphocytes

2. Test items relating to purpose

a. What is the purpose served by . . . . . ?

b. What principle is exemplified by . . . ?

c. Why is this done . . . . . ?

d. What is the most important reason for . ?

Example: Blood plasma is administered to a patient suffering from a severe burn for the purpose of

a. preventing bacterial infection
b. reducing pain
c. preventing or reducing shock
d. producing stimulation

3. Test items relating to cause

a. What is the cause of . . . . . ?

b. Under which of the following conditions is this true. . . .?

Example: What is the cause of septic inflammation?

a. Trauma or mechanical irritation
b. The effects of heat and cold
c. Chemicals from insects or plants
d. The presence of bacteria

4. Test items relating to effect

a. What is the effect of . . . . . ?

b. If this is done, what will happen . . . . . ?

c. Which of the following should be done (to achieve a given purpose. . . . . ?

Example: What will happen if the electrolyte level of a battery is allowed to remain below the tops of the plates?

a. The battery voltage will be reduced
b. The plates will expand
c. The battery capacity will be reduced
d. The sediment will rise to the top

5. Test items relating to association

a. What tends to occur in connection (temporal, causal, or concomitant association) with . . . . . ?

Example: If the metal diaphragm in the air regulator on a dental operating unit is cracked, state the position in which the valve will remain?

a. Partially open
b. Open
c. Halfway closed
d. Closed

6. Test items relating to recognition of error

a. Which of the following constitutes an error (with respect to a given situation). . . . . ?

Example: What liquid should NEVER be used for cleaning ball or roller bearings?

a. Alcohol
b. Mineral oil
c. Lard oil
d. Vegetable oil

7. Test items relating to identification of error

a. What kind of an error is this . . . . . .?

b. What is the name of this error. . . . . .?

c. What recognized principle is violated . .?

Example: While a boat is being refueled, a static spark ignited the gasoline fumes and a serious fire resulted. The man in charge of the operation might have prevented the fire if he had made certain that the nozzle of the filling hose had a

a. hard rubber surface
b. grounded wire attached
c. bakelite tip
d. 40-mesh wire gage protector

8. Test items relating to evaluation

a. What is the best evaluation of . . .(for a given purpose) and for what reason. . . . ?

Example: When the number of cases is small (e.g. less than 20) and the magnitude of the values is likewise small, the use of an assumed mean in the computation of the mean can best be evaluated as

a. less efficient than computation from actual values

b. likely to distort the value obtained by the introduction of a constant error

c. more accurate than the actual values

d. neither better nor worse than computation by other methods

9. Test items relating to difference

a. What is the (or an) important difference between a plain milling cutter and a spiral milling cutter?

a. provides a smoother cutting action
b. provides for more chip clearance
c. provides a side thrust to relieve the load on the bearing
d. make grinding easier

10. Test items relating to similarity

a. What is the (or an) important similarity between . . . ?

Example: The directional gyro and the artificial horizon utilize a

a. rigidity in space of a gyro
b. precession of a gyro
c. high speed of a gyro
d. dynamic stability of a gyro

11. Test items relating to arrangement

a. In the proper order (to achieve a given purpose or to follow a given rule), which of the following comes first (or last, or follows a given item) . . . ?

Example: In the event of a casualty to the lube oil heater, the first step to be taken is to

a. recirculate the oil
b. secure the lube oil pump
c. cut in the lube oil storage tank
d. secure the lube oil heater

12. Test items relating to incomplete arrangements

a. Before starting a grinding operation on the surface grinder, the worker should make a last check to see that

a. the wheel is properly mounted
b. the wheel is properly dressed
c. the work is properly aligned
d. the magnetic switch is turned on

13. Test items relating to a common principle

a. All except one of the following items - micrometer, vernier, caliper, gear tooth caliper, hermophrodite caliper - are properly classified as numerical measuring instruments. From the following, the one which can be substituted in the above series is:

a. depth micrometer
b. sine bar
d. surface gage
e. dividers

14. Test items relating to controversial subjects

Although not everyone agreed with the desirability to change the national holidays for our national heroes from their birthdate to a specific Monday during the month of their birth, those who supported this change did so primarily for the reason that this arrangement;

a. is better for business
b. eliminates absenteeism
c. provides an extended weekend
d. is good for public morale

15. Test items requiring the most inclusive answer

Example: Which of the following terms is the most inclusive.....?

a. smallpox
b. tuberculosis
c. scurvy
d. disease

16. Test items requiring the most dissimilar answer

Example: Which of the following terms does not belong with the other three.....?

a. shark
b. ape
c. giraffe
d. kangaroo

V. TEST ITEMS IN A NEGATIVE FORM

A. Test Items with Several Responses

A negative form is used for test items that normally have several responses which are equally good. If the problem was stated in a positive form, the alternatives for such an item would include several correct responses and one response which would not be correct or would be weaker than the others. Since it is not practical to have an item with more than one correct response, the item is prepared in a negative form. The trainee is, thus, asked to select the response that could not have been correct or would not have belonged with the other correct responses, if the item had been asked in the positive form.

1. Test items relating to the negative variety

Example

Positive statement	Negative statement
Which of the following engine lathe accessories operates on the principle of circumferential pressure:	Which of the following engine lathe accessories does NOT operate on the principle of circumferential pressure:
a. three-jaw chuck b. four-jaw chuck c. spring collet d. face plate	a. three-jaw chuck b. four-jaw chuck c. spring collet d. face plate

Comment: The item on the left in positive form has three correct responses and, for this reason, can not be used. When stated negatively, as in the item on the right, there is only one correct response.

VI. DEVELOPING MULTIPLE-CHOICE ITEMS

A. Rules for Formulating Test Items

During recent years, individuals engaged in the construction of good multiple-choice test items have developed a series of suggestions or rules that have become widely accepted. These suggestions and rules will be presented in four groups:

1. General rules
2. Rules for developing the stem or statement of the problem
3. Rules for developing the alternatives or responses
4. Miscellaneous rules and policies

B General Rules

1. The item should be expressed as clearly as possible. Unlike ordinary reading material in which extensive context helps to clarify the meaning of any particular phrase or sentence, the multiple-choice item must be clear in and of itself. The power to discriminate between the better and poorer craftsman, at a particular level, may be seriously limited by lack of clarity. In achievement testing, the difficulty of an item should relate to the problem involved rather than from the words in which it is expressed. Achievement test items should not be verbal puzzles; they should indicate whether or not students can produce the correct answer to a problem.

   Lack of clarity in a test item may arise from many causes. It may be due to the inappropriate choice and awkward arrangement of words or disorganized thinking of the test writer. Many ideas for test items are vague and general when first written. Before emerging in

final form, they need to go through a critical examination, evaluation, and revision. By doing this, clarification of ideas will go along hand-in-hand with improvement in the wording of a test item.

2. The test item, as a whole, should present a problem which, if answered correctly, will show that the student has achieved or learned an important aspect of the course of study in the subject matter area being tested.

3. The test item should be realistic and practical. It should call for knowledge with which the tradesman or industrial technician should be familiar. Asking a tradesman or technician about some trivial, insignificant, unimportant fact does not give him an opportunity to demonstrate his knowledge of the more important aspects of the subject matter areas covered by the examination. Problems which can be answered by general intelligence or general knowledge should be avoided.

a. Test item relating to practical problems

Examples

Not a practical problem

Which of the following is the formula to use in determining board feet?

a. $\frac{T' \times W' \times L''}{144}$

b. $\frac{T'' \times W'' \times L'}{12}$

c. $\frac{T'' \times W'' \times L''}{144}$

d. $\frac{T'' \times W' \times L''}{12}$

Practical problem

In order to construct a frame, two pieces of cypress 1" thick, 8" wide, and 6' long are needed. If the cost of cypress is 24¢ per board foot, what will the lumber cost?

Comment: In the item asking for the board foot formula, the student is very apt to say, "So what?". It is doubtful that he would have to recite the formula for finding board feet. That is all the item would measure. It is not too practical. If it is desired to measure the ability of the individual to use the formula, this can be accomplished much better with the more practical problem which deals with information that may have to be used or known on the job.

b. Test items relating to useful information

Examples

Useless information	Useful information
How many basic joints are used in welding?	It is not advisable to attempt to weld cast iron to steel with a steel filler rod because
a. five	a. cast iron has a higher melting point than steel
b. two	b. steel has a higher melting point than cast iron
c. three	c. neither metal is stable at welding temperatures
d. nine	d. the iron will ruin the structure of the steel molecules

Comment: The information required to answer the item at the right correctly is useful on the job. There is no point in asking the item at the left, since a knowledge of how many joints are used in welding is of no value as far as doing a welding job is concerned.

c. Test items involving general versus specific information

Examples

General knowledge	Specific information
In arc welding the greatest hazard is	To guard against the greatest safety hazard in arc welding it is essential at all times to
a. burns from hot sparks and slag	a. wear a protective apron
b. burn damage to the feet	b. wear safety shoes
c. injury to the eyes	c. wear a helmet with proper lens
d. flash burns	d. ground the work properly

Comment: Both items are supposed to test knowledge of the most serious hazard in arc welding. The answer on the left can easily be guessed. More specific knowledge is required to select the correct item on the right.

d. Test items involving appropriate-inappropriate statements

The problem should be stated in language which is common or appropriate to the job and/or subject matter. It is important to avoid copying a problem directly from a manual or a textbook. Some examples of statements of problems inappropriately and appropriately worded follow.

Examples

Inappropriate statements	More appropriate statements
1. What is the function of the clapper box on a shaper?	1. The clapper box on a shaper is used for .....
2. The main factor which determines the speed of operation of the acetylene welder is the	2. The speed of work of the acetylene welder depends mainly on the .....
3. What operation must be undertaken first in sharpening a handsaw?	3. The first thing to do in sharpening a handsaw is to .....

e. Test items involving textbook or job language

Two other examples, taken from achievement tests, are given below:

Examples

Textbook language	Job language
Which of the following oscillators is employed to take advantage of the interelectrode capacity of a triode?	Which of the following oscillators makes use of the interelectrode capacity of a triode?
a. tuned plate - tuned grid b. relaxation c. Hartley d. colpit	a. tuned plate - tuned grid b. relaxation c. Hartley d. colpit

Comment: The item at the left was copied from a manual. It emphasized a relatively unimportant point, authorization. The language is awkward and tends to make the item difficult to read. The item at the right is worded in language which the trainee will use on the job and should understand.

f. Test items involving clarity of words

Persons responsible for writing items should choose words that have precise meaning whenever possible. Lack of clarity frequently arises from inappropriate word choices. Many commonly used words and phrases have no precise meaning. Others have no meaning that applies accurately in the context in which they appear.

Example: In the 1948 presidential campaign, Truman yielded unofficially to Dewey on which of the following points?

a. That the un-American activities investigations were not direct attacks on the Democratic Party

b. That the tone of the political campaign would be kept on a high level

c. That the civil rights program be abandoned

d. That there should be an equal balance of Democrats and Republicans in Congress to insure sound legislation

Comment: The words "yielded", "unofficially", and "points" are vague.

g. Test items involving simple sentences

Personnel writing items should avoid complex or awkward arrangements. In regard to this, F. L. Ebel makes the following suggestions:

"The structure of sentences used should be as simple as possible. It is often advantageous to break up a complex sentence into two or more separate sentences. A qualifying phrase should be placed near the term it qualifies. In general, it is desirable to make clear the focal point of the problem early in the sentence and add qualifications or explanations later. Finally, it is often helpful for the item writer to ask himself, 'Just what is the point of this item?'. In the answer to this question, an item writer may find a simpler, more direct wording; or, he may discover that it has little point; or one that is not worth testing."

h. Test items must be independent of each other

The items in an examination should be independent of each other. The information contained in one item should not give away the answer to another item.

Examples

When a naval vessel is at anchor or moored, it flies the National Ensign at the

a. mainmast
b. foremast
c. jackstaff
d. flagstaff

When a ship is not underway, the National Ensign should be displayed from the flagstaff from

a. 0800 to sunset
b. sunrise to 1830
c. 0800 to 1830
d. sunrise to sunset

Comment: It is obvious that the statement of the problem in the item at the right gives away the answer to the item at the left.

4. Judging the Difficulty Level of Questions

The level of difficulty of the item should be adapted to the group and purpose for which the item is intended. While subjective judgements of item difficulty are not too accurate, personnel who are well acquainted with the general ability and knowledge of the competency level and who know something about typical performances on similar and old items can do a useful job of judging item difficulty.

5. Pitfalls in Judging Item Difficulties

In connection with the adjustment of item difficulty, two pitfalls must be avoided. The first is the application of a "minimum essentials" concept to an achievement test. According to this concept, a test should include only those items that all journeymen should, according to minimum essential standards, be able to answer correctly. A test consisting of items meeting this requirement will, in all probability, be too easy.

The second pitfall (or mistake) is writing (or selecting) items from the standpoint of what the ideal journeyman should know rather than in terms of what the typical tradesman knows. A person who is unrealistic about the typical tradesman is likely to construct or select unreasonably difficult test items. Some people believe that such a test writer has high standards. It would be more appropriate to say that such a person is unrealistic.

a. Test items involving increasing difficulties

It is desirable, and often necessary, to sample an area of knowledge at different levels of difficulty. The following example illustrates how items of increasing difficulty can be written to cover the same subject matter.

Example

Easy: Which of the following should be used to extinguish a fire in a small electric motor?

a. A bucket of sand
b. A stream of water
c. A portable fire extinguisher
d. A blanket or a mattress

More difficult: What type of fire extinguisher is the best for extinguishing a fire in an electric motor?

a. Foam
b. CO_2
c. Pyrene
d. Carbon Tetrachloride

More difficult: Even though a pyrene extinguisher is almost as effective as a CO_2 extinguisher for extinguishing a fire in an electric motor, it should NEVER be used due to the fact pyrene will

a. form a toxic gas when contacting hot metal
b. conduct electric current
c. render motor unfit for repair or salvage
d. explode when near an electric spark

Most difficult: What is the primary extinguishing action that takes place when a CO_2 extinguisher is used on a fire in an electric motor?

a. The heat of the fire causes the CO_2 to expand which thins out the oxygen below the amount needed for combustion.

b. The CO_2 being extremely cold, quickly cools every part of the motor below its flash point.

c. The CO_2 because of its great density, settles in a blanket, displacing the air and oxygen necessary for combustion.

d. The CO_2 combines chemically with the air to form CN_2O_3 which effectively removes oxygen from the fire area.

VII.--THE CONTROL OF DIFFICULTY LEVEL FOR TEST ITEMS

A. Ways of Changing the Difficulty of a Test Item

There are two ways of changing the difficulty of a test item: (1) change the statement of the problem, or (2) change the alternatives. Changing item difficulty by changing the alternatives is quite limited in the practical situation. The statement of the problem usually determines the range of possible alternatives. Alternatives which are altered toward increased complexity and, hence, difficulty, would probably necessitate a more difficult or complex statement of the problem.

It should be apparent, then, that changing item difficulty by changing the statement of the problem is probably the better approach. It is much easier and more satisfactory to change the statement of the problem and find suitable alternatives, especially the distractors, to the statement of the problem. It is important to remember that distractors should do just what their name implies. They should distract the less qualified students from the correct response. Distractors can do this only if they are plausible answers Items sometimes appear easier because distractors fail to distract.

B. Check Points for Evaluating Questions

Can you answer "yes" to each of the following questions about each of the test items you have written?

Is the test item:

1. stated as clearly as possible,
2. realistic and practical,
3. dealing with an important aspect and knowledge of the job,
4. phrased in the language of the job,
5. stated in words having precise meaning,
6. presenting an independent problem,
7. of appropriate difficulty?

If you can answer "yes" to each of the above questions concerning test items which you have written or revised, then your items should be specific in terms of content and purpose.

VIII. RULES FOR DEVELOPING THE STEM OF THE TEST ITEM

A. Test Items Involving a Clear Central Problem

The statement of the problem should be or contain a clear central issue to be solved. If a well-informed student can give the correct answer after reading the statement of the problem (which in this case is written as an incomplete statement) without reading the alternatives, it can be assumed that a clear central problem has been stated.

Examples

No Central Problem	Clear Central Problem
The persistency of a gas	The persistency of a gas refers to its
a. refers to its toxic effects	a. toxic effects
b. refers to its identifying odor	b. identifying color
c. is the duration of its effectiveness after release	c. duration of effectiveness after release
d. is its clinging effect on the skin	d. clinging effect on skin

Comments: The stem of the problem at the left (the persistency of a gas) is really not a problem at all because it has no central thought. The candidate does not know, from the stem, what problem is to be solved. The item can be improved by bringing the central problem into the item stem as illustrated at the right.

B. Test Items Involving Pure Memory vs. Understanding and Principles

When developing the statements of the problem try, as much as possible, to avoid a problem which requires pure memory for its solution. A candidate may know words, or phrases, without knowing their meaning or application. Items involving only memory for solutions will probably not tell which of the candidates have gained the greatest mastery of the subject matter. Items which call for an understanding and application of facts and principles are those most apt to tell what needs to be known about each candidate's competency. Candidates who are most successful on the job as craftsmen are those who are able to apply to job situations facts and principles which they have learned.

C. Test Items Must Be Stated Briefly and Completely

The problem should be stated clearly, briefly, and completely. It should be stated in the fewest possible words, but it should include all the essential information in relation to the alternatives.

Example

Lack of information	Complete information
An external thread would be	A hardened, threaded section of a work piece is machined to a precise micro-finish by
a. turned	a. turning on a lathe
b. cut with a die	b. cutting with a screw threading die
c. ground	c. grinding or using a precision grinder attachment
d. tapped	d. tapping

Comment: The item at the left does not give sufficient information upon which to base an intelligent answer. There is no problem and there is no single answer as there are the three underlined answers. Additional information in the item at the right makes the problem clear and a single answer possible.

D. The Problem Should Contain Only Material Applicable to its Solution.

Information which is not necessary for obtaining an answer to the problem should be omitted. Such information may tend to destroy the validity of the item as a measure of what it is intended to measure. Since students vary greatly in reading ability, it is undesirable for reading ability to be a factor in determining the correct response. In general, items should be kept as short as is consistent with clear statements. Some introductory statements that are not strictly necessary may occasionally be justified if they help to clarify the problem or to establish its importance.

Example

Irrelevant information	Relevant information
The electronic serviceman must be able to identify certain causes of difficulties in receivers by their sound. If you hear a high pitched squeal in the output of an AM receiver, this defect is usually caused by	A high pitched squeal in the output of an AM receiver is usually caused by
a. excessive amplification b. poor tracking of the local oscillator c. a defective output transformer d. a defective detector	a. excessive amplification b. poor tracking of the local oscillator c. a defective output transformer d. a defective detector

Comment: The information contained in the item at the left is not necessarily to present the problem. The irrelevant information requires additional reading time of the trainee and makes it more difficult for him to determine the nature of the problem.

E. The Ability to Select Material From the Problem Which is Applicable to the Correct Solution is Often Quite Important

Example

What is the amount of net increase inventory in the following situation?

Closing inventory	$ 91,000
Purchases	$ 80,000
Sales	$ 62,000
Opening inventory	$ 58,000

a. $149,000
b. $ 51,000
c. $ 33,000
d. $ 15,000

Comment: In the preceding problem, a person must not only be able to perform the arithmetic involved but he must be able to select the relevant information from the problem. Irrelevant information in the problem (purchases and sales), which is common in a practical solution, is necessary in order to test the trainee's ability to select the information he needs to arrive at the correct answer.

F. The Problem Should be Stated in Positive Form if Possible

Students may respond to an item in the negative form as if it were stated in the positive form. Capitalizing words such as NOT, POOREST, NO, LEAST, NEVER, etc. may help to overcome this to some extent. If these words, however, appear in the statement of the problem and they neither add nor detract from the desired emphasis, they should, naturally, be in the lower case. A problem usually has more significance and is a better problem if it is stated positively

G. When Writing Reasoning Problems in Related Mathematics, Science, Drawing Interpretation, Avoid Making the Computational Parts of the Problem Unintentionally Difficult

Trainees frequently answer reasoning problems in mathematics, science and other related subjects incorrectly, even though they have reasoned correctly. This is due to a slight error in computation, and not in reasoning. The person writing the item should decide whether he wishes to test computational skill involving complex or time-consuming calculations, or an understanding of mathematical principles. In this latter case, the test item should keep computational difficulty to a minimum.

H. The Stem Should Include All Qualifying Information Needed to Provide a Reasonable Basis for Response Selection

Sometimes, a person writing an item does not state explicity all of the qualifying information. A person taking an examination must have all qualifications specifically stated.

Example

If a ship is wrecked in very deep water, how far will it sink?

a. just under the surface
b. to the bottom
c. to a depth where the pressure is equal to its weight
d. to a depth which depends in part upon the amount of air remaining in the ship

Comment: A number of capable people would probably select response number d, instead of the desired correct response b, because they considered the possibility (which the person that wrote the item failed to exclude) that a wrecked ship might not sink completely but might remain partly submerged. In that case, response d, while not too good, is the best response available.

I. When a Stem Contains Lengthy or Involved Qualifying Information, It is Usually Clearer to State the Qualifying Information First and End the Statement of the Problem Requesting the Desired Information.

It is, sometimes, clearer to state the qualifying information first as a complete sentence, and then use another sentence to state the problem, or implied question. The following suggestions may be helpful when developing an item stem:

1. If there is no qualifying information, state the question first (who, what, where, when, etc.).

2. If there is one qualifying clause, state the question last if this makes the problem clearer.

3. If there are two or more qualifying clauses, use more than one sentence and state the question in the last sentence.

It is well to remember that clarity is the main consideration.

Clarity of Statement:

Qualifying information last	Qualifying information first
What safety precautions must be observed first when dressing or finishing the edge of a board on the jointer?	When dressing or finishing the edge of a board on a jointer, what safety precautions should be observed first?

Comment: When the qualifying information is stated first, as in the stem on the right, the situation upon which the question is based is known to the person taking the test, sooner than in the stem at the left. When the necessary information is identified first (before the question is asked), the entire problem is clearer and the requirements are more quickly grasped.

IX. RULES FOR DEVELOPING THE ALTERNATIVES

A. The Correct Response Must Be Right

The correct response, or the best response, should be unquestionably right. The reference from which the information for the item is taken should be currently accepted as authoritative. In order to avoid confusion, a partially correct response should not be used even if it is the best of the given alternatives. This can, frequently, be remedied by rewriting the statement of the problem. Instead of asking, for example, "What was the purpose of .....?"; rewrite it to ask, "One purpose of ..... was to .....

B. The Distractors (wrong responses) Should Be Representative of Errors Which Are Commonly Made By Less Qualified Tradesmen

Such errors are usually due to errors in technical judgement. These errors may be used as distractors which discriminate between the better qualified and the less qualified candidates. Popular misconceptions held by people, in general, should not be used as distractors. Such misconceptions are usually due to ignorance and should have been corrected in the teaching/learning situation.

C. Each of the Alternatives Should Literally Be a Possible Answer

The distractors should be closely related to the problem to attract the poorer trainees, but not any large percent of the better qualified trainees to use these as answers. The distractors should be important, plausible answers. They should not be obvious errors.

Example

Unrelated wrong answers

One of the requirements for advancement in grade in the Civil Service is that the candidate must

a. be working in the department
b. be married and have a family
c. pass a qualifying examination
d. be recommended by his superior

Plausible wrong answers

One of the requirements for advancement in grade in the Civil Service is that the candidate must

a. have at least three months experience in the higher grade
b. be recommended by the personnel office
c. pass a qualifying examination
d. have at least two years experience in the work of the higher grade

Comment: The distractors, as plausible wrong answers, are much more likely to be used by the person who does not know anything about the subject than the unrelated distractors which are obviously wrong.

D. All of the Distractors Should Be Appropriate to the Item Stem

Example

Why do living organisms need oxygen?

a. Purification of the blood
b. Oxidation
c. Release of energy
d. Assimilation of the blood

87

Comment: In this item the alternatives are not appropriate to the item stem for the responses are not stated as reasons which are required by the stem.

E. The Item Writers Should Avoid Irrelevant Clues Which May Give Away the Correct Response

If a number of candidates who would not normally be able to choose the correct response notice the clue and respond correctly, the item is definitely weakened in that it measures something other than that for which it was intended.

1. Irrelevant clues exist whenever there are any systematic formal differences between the correct response and the distractors.

A good example would be when there is a tendency to place the answer in one favored position among the several alternatives. If a trainee should observe that the third response is most frequently the correct answer, or that the fourth response is seldom correct, he may use such observations as a foundation for successful guesses on other items. Test developers, responsible for writing items or for assembling examinations, should be particularly alert not to favor any one number as the correct response, and to attempt to get a good distribution among all the numbers for correct response.

2. Probably one of the most common types of irrelevant clues involves common elements in the item stem and in the answer.

An example of an item having common elements and a reworked or corrected item is given.

Example

Common element

The divisor is a term used in

a. addition
b. subtraction
c. division
d. multiplication

No clue

What arithmetical method should be used to find the number of square feet of floor space in a room which is 12 feet wide and 24 feet long?

a. division
b. multiplication
c. subtraction
d. addition

Comment: In the item at the left there is an element which is common to both the problem and the correct response. The word "divisor" in the problem gives away the answer "division" in the alternatives. This is so obvious that most any untrained person could get the correct answer. When there is no clue, a knowledge of arithmetical methods is necessary to find the correct answer.

3. Words like "all", "more", "certainly", "never", and "always" are called specific determiners.

The use of specific determiners tends to provide irrelevent clues to the correct response because statements including them (unless great care is taken) are predominantly false. Similarly, modifiers such as "sometimes", "generally", and "usually" should be used in both correct and incorrect alternatives, if they are used at all, since they, otherwise, may give a clue to the correct answer.

4. The length of the correct response should not be a clue to the correct response.

If the correct response is consistently either longer or shorter than the other alternatives, it is possible that the trainees will discover this and rely on this tendency rather than on knowledge of the subject matter in the examination.

5. Clues due to grammatically incorrect sentences.

A clue to the correct response or to the elimination of some distractors is sometimes due to the fact that the statement of the problem does not make a grammatically correct sentence in combination with each alternative. This should be avoided.

Examples

Grammatically incorrect	Grammatically correct
On a road map, a main highway is shown by a	On a road map, a main highway is shown by a
a. single heavy line	a. single heavy line
b. two red dotted lines	b. pair of dotted lines
c. parallel broken lines	c. set of parallel broken lines
d. parallel unbroken lines	d. set of parallel unbroken lines

Another example which is very common follows:

When filing metal, the file should be cleaned frequently with a	When filing metal, the file should be cleaned frequently with a/an
a. wire brush	a. wire brush
b. oily rag	b. oil rag
c. emery board	c. emery board
d. file card	d. file card

Comment: In the item at the left, it should be obvious that since alternatives (b) and (c) make the problem grammatically incorrect, that there are only two possible answers, (a) and (d). This situation is remedied, in the

item at the right, by adding "an" along with "a" at the end of the problem. Each alternative on the right makes a grammatically correct sentence and the possibility of guessing the correct answer is reduced.

6. Alternatives should deal with similar materials or elements

The alternatives should, usually, deal with similar ideas, materials, or elements which are expressed in parallel form.

Example

Unrelated alternatives	Similar alternatives
In the Civil Servica a beginning typist must:	According to the statement of Qualifications of the Civil Service a candidate for beginning typist must be able to type from copy at the rate of, at least, how many net words per minute?
a. be able to type from copy at the rate of 50 net words per minute	a. 30
b. set up and maintain confidential files	b. 35
c. dust the desk	c. 40
d. be punctual	d. 50

Comment: In the item at the left, the lack of parallel structure and similarity of material confuses the whole problem and provides a clue that alternative (a) was probably intended for the best answer. The item at the right deals with similar material and the problem is clear. Each alternative is in terms of net words per minute.

7. Irrelevant clues may be constructive on rare occasions

It should be noted that irrelevant clues may, on rare occasions, be used in a constructive manner. The deliberate planting of some of the clues mentioned above may tend to defeat the rote-learner or to make the distractors highly attractive to those trainees whose knowledge is superficial.

F. Each Response Must Make Grammatical Sense

Alternatives are, frequently, made unnecessarily long by including in each alternative material which should be included in the stem. The stem should, in general, include any words that would otherwise have to be repeated in each response. It should be noted, however, that it is not always desirable to eliminate all words common to the responses. Sometimes, it is necessary to introduce each response with a word or phrase to make grammatical sense. Also, if the retention of common words in all the responses makes the item easier to understand, they may be retained. Usually, however, the item writer will want to transfer common words to the stem.

Example

Material repeated in each alternative	Material taken from alternatives and placed in the problem
A portable electric submersible pump with a clean strainer that is pumping against a 50-foot head or less will discharge.	How many gallons of water per minute will be discharged by a portable electric submersible pump, with a clean strainer, that is pumping against a 50-foot head or less?
a. 100 gallons of water per minute b. 140 gallons of water per minute c. 160 gallons of water per minute d. 180 gallons of water per minute	a. 100 c. 160 b. 140 d. 180

Comment: The problem at the right is more easily read and understood than the one at the left because material which is a part of the problem is not repeated in the alternatives.

G. Responses That Overlap Or Include Each Other Should Be Avoided

Example

What percent of the total loss, due to hail, is the loss to growing crops?

a. Less than 20%
b. Less than 25%
c. More than 50%
d. More than 55%

Comment: This item is actually a two-response item. The choice is between (a) and (c) for, if (a) is correct, then (b) is also correct; if (d) is correct, (c) is also correct.

H. Acceptable Alternatives

1. "None of the above", in some items, is an acceptable alternative. For example, the subject of safety precautions is an area in which it is sometimes necessary to find out whether or not the individual would not, under any circumstances, commit the act presented in any of the alternatives. Another use for "None of the above" is as a distractor when the correct response seems to give itself away and would be too easy to see in a set of ordinary alternatives.

"None of the above" is sometimes used in mathematical problems as either the correct response or as a distractor. In all cases, it is used as the last alternative. Some people feel that "None of the above" should be used as a response only in items which are of the best--answer type. There are a few limitations in connection with using it as an alternative. It SHOULD NOT be used if the statement of the problem is an incomplete statement.

A pattern should not be developed in using it, and it should be used about as frequently as a correct response as it is as an incorrect response. It should not be used just to "fill in" when a final distractor is difficult to develop.

2. "All of the above", in some items, is an acceptable alternative. There is some difference of opinion, however, concerning its proper use. The same discretion should be followed when using "All of the above" as with the use of "None of the above". In all instances, it must be used as the last alternative. It, too, should not be used as an alternative following the statement of a problem which is an incomplete statement.

3. "Any of the above" will not be used as an alternative.

4. Items requiring negative alternatives should be used as little as possible. Items with positive alternatives may be clearer to the candidates. The nature of the subject matter usually determines the use of this type of alternative. The following examples are given to illustrate this type of item:

Examples

Negative alternatives

When refilling water glasses during a meal, you should

a. remove the glass from the table
b. always place the glass on a serving tray
c. not move the glass from its position on the table
d. never move the glass more than six inches to the right

Positive alternatives

What should you do with a water glass when refilling it during a meal?

a. remove it from the table
b. place it on a serving tray
c. leave it in its position on the table
d. move it six inches to the right

Comment: In the item at the left, alternatives (c) and (d) are stated negatively. In the item at the right, all of the alternatives are stated positively.

5. Instead of repeating two alternatives in combination to form a third alternative, variety is added to the item by using the alternative numbers instead of repeating the alternatives themselves.

Examples

Repeating alternatives

What measure(s) of central tendency should be used when determining the score which occurred most often among a set of examination scores?

a. Mean
b. Mode
c. Mean and mode
d. Median

What will most likely cause an individual to obtain a different score on the same test when taking it a second time?

a. Practice effect
b. Additional study
c. Practice effect and additional study
d. None of the above

Using alternative numbers

What measure(s) of central tendency should be used when determining the score which occurred most frequently among a set of examination scores?

a. Mean
b. Mode
c. Both (a) and (b) above
d. Median

What will most likely cause an individual to obtain a different score on the same test when taking it a second time?

a. Practice effect
b. Additional study
c. Both (a) and (b) above
d. None of the above

X. MISCELLANEOUS RULES AND POLICIES

A. Suggestions Pertaining to Numbers

The following suggestions pertain to the handling of numbers in the writing of multiple-choice items:

1. It is suggested that the numbers from one to ten be spelled out, except in the following situations:

   a. In mathematical problems, it is common practice to use numerals (figures).

   b. When a number greater than ten appears in the statement of the problem, or question, or in the alternatives, or both, all of the numbers used in the item should appear as numerals.

   c. Fractions appearing in mathematical problems, which are accompanied by a whole number, should appear as mixed numbers regardless of whether or not the whole number is less than ten.

   d. Decimals should always appear as Arabic numerals.

2. Fractions, appearing alone as alternatives, should be arranged neatly in one of the following patterns:

a. 1/2	a. 1/2	a. 2 1/4
b. 3/4	b. 3 1/4	b. 8 1/4
c. 5/6	c. 9 1/2	c. 76 1/4
d. 7/8	d. 9 3/4	d. 108 1/4

3. Numbers should appear in ascending or descending order.

Example

Descending order

a. 40	c. 20
b. 30	d. 10

Ascending order

a. 10	c. 30
b. 20	d. 40

4. When the numbers one through four stand alone as alternatives, or when the numbers 1, 2, 3, or 4 appear as alternatives, they should be placed beside their corresponding digit as follows:

a. One	a. 1
b. Two	b. 2
c. Three	c. 3
d. Four	d. 4

5. When decimals are used, the decimal points should be in line.

Example

a. 12.34
b. 123.45
c. 1,234.56
d. 12,345.67

6. Monetary amounts should appear as follows:

a. $0.25	a. $ 5.50	a. $ 2.50
b. $0.55	b. $ 50.50	b. $ 20.00
c. $0.65	c. $ 75.15	c. $ 200.50
d. $1.50	d. $100.00	d. $2000.00

B. Rules Pertaining to Capitalization

The following rules, pertaining to capitalization, should be observed in the preparation of test items:

1. The first letter of the first word of each problem should be capitalized.

2. Whenever the stem of an item is in the form of a direct question, the first letter of the first word of each alternative should be capitalized.

3. Whenever the stem of an item is in the form of an incomplete statement, the first letter of the first word of each alternative should not be capitalized, unless it is a proper name and would ordinarily be capitalized regardless of where its position is in the alternative.

4. The first letter of all proper names and abbreviations of proper names should be capitalized.

C. Punctuation After Alternatives

No punctuation is necessary after ordinary alternatives. This refers to periods after each alternative

XI. VARIATIONS IN MULTIPLE-CHOICE TEST ITEMS

A. Test Items Involving Multiple Completion

Of the three or more possible responses offered in an item of this type, one or more may properly complete the stem (premise). The item is answered correctly only when the appropriate response or combination of responses has been selected. Thus, the examinee can reach a correct answer only by committing himself as to the relevance, or irrelevance, of each one of the three or more possible choices. This selection process requires judgment in relating various considerations to the whole, and depends, especially, upon the completeness of the candidate's knowledge and experience with a particular problem. Obviously, use of this type should be confined to situations where several considerations of equal importance and plausibility can be suggested and are representative of the kind of decision which the craftsman must make in the occupation.

B. Application of Multiple Completion Test Item

Excellent use can be made of the multiple completion technique wherever a problem involves:

1. a situation which may require a consideration of several possible consequences.

Example

Using inert electrodes, direct current is passed through a concentrated water solution of sodium chloride. This would result in the

a. evolution of chlorine gas
b. formation of sodium hydroxide solution
c. plating out of metallic sodium on the cathode

(a, b) correct

2. a situation where one or more conditions must be specified in order to define it adequately,

Example

Isotopes of a given element are related in that they have the same

a. number of neutrons in the nucleus
b. position in the periodic table
c. arrangement of planetary electrons

(b, c) correct

3. two situations which are alike (or unlike) with regard to one or more points of comparison,

Example

In both meiosis and mitosis, there normally occurs

a. polar body formation
b. homologous chromosome pairing
c. cleavage of the cytoplasm

(c) correct

4. a principle which may be applicable to one or more situations,

Example

Joules are units by which one can express

a. work
b. energy
c. momentum

(a, b) correct

5. several considerations which may bear upon a single result.

Example

The rate of a chemical reaction may be influenced by

a. temperature
b. catalysts
c. pressure

(a, b, c) correct

C. Appearance (Structure) of the Multiple Completion Test Item Type

Directions: For each of the test items (1 through 5), ONE or MORE of the responses given are correct. Decide which response(s) is (are) correct and on the answer sheet blacken the correct space.

1. if a, b, and c are correct;

2. if only a and b are correct;

3. if only b and c are correct;

4. if only a is correct;

5. if only c is correct;

MARK ONE SPACE ONLY ON YOUR ANSWER SHEET FOR EACH QUESTION

Figure 1. Directions Summarized

Blacken Space	1	2	3	4	5
For Correct Responses	a, b, c	a, b only	b, c only	a only	c only

D. Hints in the Construction of a Multiple Completion Item

1. Each of the choices should be able to stand alone without depending upon (or excluding) any other choice.

Example

In studying the rate of oxygen production by photosynthesis during a laboratory experiment, it is necessary to

Good	Poor
a. regulate the flow of CO_2	a. generate CO_2
b. control the light intensity	b. provide a flow of CO_2
c. maintain constant temperature	c. measure the flow of CO_2

2. Make the choices refer to as many as possible of the pertinent factors which relate to a problem. (A choice can occasionally be stated falsely to serve as a distractor.)

Example

The administration of adrenaline may lead to

Good	Poor
a. bronchiole contraction	a. increased blood pressure
b. increased blood pressure	b. a harder working heart
c. increased blood sugar	c. increased photosynthesis

(b, c) correct

3. In the stem avoid giving any clue as to how many of the choices may be correct.

Example

GOOD: Hybrid sterility may be caused by ..... (followed by multiple responses)

POOR: The genetic factors in hybrid sterility are ..... (followed by multiple responses)

4. When submitting multiple completion items, each correct choice should be designated.

5. Never have an item with no correct choice; the directions assert that at least one choice must be correct.

XII. CLASSIFICATION ITEMS

A. Test Items with a Number of Key Topics

The classification item-type presents key topics, usually concepts or principles, followed by a set of phrases. Each phrase is associated with, or related to, one of the topics. This item-type provides a quick means (in terms of testing time) of determining the candidate's mastery of the topics presented. The degree of mastery required of the candidate is very much dependent upon the skill manifested by the item-writer in framing the set. The set may merely probe familiarity with the concept or principle present. On the other hand, the set may probe the candidate's ability to apply the concept or principle to a novel situation.

It is quite important that the set be so constructed that the key topics have a somewhat reasonable relationship to each other in order to insure that the examinee is required to exercise an appropriate amount of discrimination.

1. Examples of Classification Items

Directions: Each set of test items consists of five laws or principles followed by a list of words or phrases which are numbered. For each law or principle select the one principle which is most closely related to it. Blacken the corresponding lettered space on the answer sheet to indicate the correct corresponding law or principle. One law or principle may be used once, more than once, or not at all.

Set I

Law or Principle	Condition
(A) Boyle's law	1. Is fundamental to the operation of a spray atomizer
(B) Charles' law (Gay-Lussac's law)	2. Accounts for the apparent loss of weight by an object when it is submerged
(C) Archimedes' principle	3. Explains the operation of hydraulic brakes
(D) Pascal's law	4. Explains why the canvas tops of convertible automobiles bulge upwards while the car is moving
(E) Bernoulli's principle	5. Can be used to explain why an inflated toy balloon gets bigger when heated
	6. Accounts for the fact that an air bubble gets larger as it rises to the surface of a lake

In conditions 1 through 6, the examinee's understanding of each law is tested, not by asking him to state the law, but by establishing whether he can use it to explain ordinary phenomena.

Other sets of classification test items would be included on such a test.

XIII. FIVE-RESPONSE MULTIPLE-CHOICE TEST ITEMS

A. The Test Item Type Most Widely Used in Objective Tests is the Simple Five-Response Multiple-Choice Completion Type

This item type is so familiar to most students that they handle test items with little communication difficulty. In addition to this advantage, broad experience in many types of test programs has shown this

item type to be a consistently good one in discriminating degree of achievement and ability from the less able students. In writing this type of test item, the material should be searching, demanding thought rather than the mere recognition of rote-recall answers.

B. Appropriateness of Five-Response Multiple-Choice Completion Item Type

The problem is framed in the form either of an incomplete statement or of a question. Four examples are used to show different applications.

Directions: Each of the questions, or incomplete statements, is followed by five suggested answers, or completions. Select the one which is best in each case. Blacken the corresponding lettered space on the answer sheet.

1. When the problem presented has a unique solution

Example: A cell having a flat, disc-shaped construction can exchange a given quantity of gas more quickly than a spherical cell of the same volume because a flat shape allows for

(A) a thinner cell membrane
(B) greater cell mobility
(C) greater plasticity of the protoplasm
(D) greater structural reinforcement
(E) a relatively greater surface area (Key E)

2. When the problem presented is clearly delimited by the wording of the test item so that the student selects not a universal solution but the best of the solutions offered

Example: Of the following, which is an example of a simple machine designed to multiply distance rather than force?

(A) a pair of pliers
(B) a single movable pulley
(C) a wrecking bar
(D) a cane fishing pole
(E) a nutcracker (Key D)

3. When the problem is such that the student is required to evaluate the relevance of five plausible, or even scientifically accurate, options and to select the one most pertinent.

Example: Which of the following statements would be most helpful in explaining how a lens is able to bend light rays?

(A) light reflects from the surface of glass
(B) light travels more slowly in glass than in air
(C) light striking a glass surface at a certain angle is reflected at the same angle
(D) light is a form of energy
(E) light can move through a vacuum (Key B)

4. When the problem has several pertinent solutions and the student is required to select the one inappropriate solution which is presented

Example: All of the following are factors known to upset the law of genetic equilibrium EXCEPT

(A) population size
(B) natural selection
(C) random reproduction
(D) genetic drift
(E) migration

(Key C)

C. General Suggestions For Preparing Five-Response Multiple-Choice Items

1. The test item in general

a. Each item should test a concept or idea that is important for the student to know or understand.

b. Items should vary in difficulty. For most tests, the difficulty of the "average item" should be such that 50-60% of the students get it right. In addition, there should be items which can be answered by as few as 15%, or as many as 85%. Most writers of test items tend to underestimate the difficulty of the items they produce.

c. The language of the item should be appropriate to the educational level of the students who are to answer it.

d. The wording of the item should be simple, direct, and free of ambiguity. The wording should be edited for brevity. Unnecessary words are time consuming.

e. Items containing double negatives are likely to cause confusion. If a word such as "not" or "false" appears in the stem, try to avoid using another negative word either in the stem or in any of the options.

2. The stem

The stem of the item should present clearly the central problem or idea. The function of the stem is to set the stage for the options which follow. The poor example which follows shows that if this principle is not heeded, the choices tend to become a group of not very closely related true-false questions.

Poor: An ion

(A) is any charges particle
(B) is formed when any substance dissolves in water
(C) is formed when atoms gain or lose electrons
(D) is present only in solutions which are being electrolyzed
(E) results only when atoms gain electrons

(Key C)

Good: Ions of calcium differ from atoms of calcium in all of the following ways EXCEPT

(A) their ease of reduction
(B) their ease of oxidation
(C) their reactivity with water
(D) their nuclear charge
(E) the number of extranuclear electrons (Key D)

3. The options (possible responses)

a. The options in an item are as important as the statement of the problem in the stem. They should be formulated with care; incorrectness should not be the only criterion. The following types of incorrect options should be avoided:

1. a response which is related to the situation and which sounds plausible to the untutored, but which is incorrect,

2. a common misconception,

3. a statement which, in itself, is true but does not satisfy the requirements of the problem, and

4. a statement which is either too broad or too narrow for the requirements of the problem.

b. The options should not include "specific determiners". Candidates who are "test-wise" can usually detect that statements containing words like "all", "always", "no", "never", and similar terms are generalizations so broad that they are likely to be false. Statements, on the other hand, which contain limiting words like "usually" or "sometimes" are more likely to be true. In writing options, one should try to avoid the use of "specific determiners" like these since some candidates, by logic alone, get credit for knowledge they do not possess.

c. One must avoid wording the correct response so much more carefully than the other options that it becomes conspicuously longer. In general, the correct response should be no longer than the other options.

XIV. GENERAL CHECKLIST FOR MULTIPLE-CHOICE ITEMS

A. The following checklist may be used to check on whether the suggested procedures for test item development have been observed:

1. Is the item stated as clearly as possible?

2. Is the item, as a whole, realistic and practical?

3. Does the item, as a whole, deal with an important and useful aspect of the subject matter?

4. Is each item independent of every other item in the test?

5. Is the item, as a whole, specific?

6. Is the item of appropriate difficulty?

7. Does the item have a clear central problem?

8. Is the item measuring something more than pure memory?

9. Is the item stated accurately?

10. Have the problem and alternatives been stated briefly, but completely?

11. Does the problem contain only material relevant to the solution?

12. Does the problem, or the statement, of the stem contain all the qualifying material necessary to provide a reasonable basis for selecting the correct response?

13. Are the distractors (wrong responses) important, plausible answers rather than obviously incorrect distractors?

14. Are all of the distractors appropriate to the item stem?

15. Does the item stem avoid giving away the best response by irrelevant clues?

16. Do the alternatives overlap or include each other?

17. If "None of the above", or "All of the above", have been used as alternatives, have they been used appropriately?

18. Does any item contain "Any of the above" as an alternative?

XV. LENGTH OF WRITTEN TESTS

To test for all areas of knowledge or skill would require an extremely long test. A well-designed test contains a sufficient number of test items to show the degree to which tradesmen and technicians possess the knowledge and skills in comparison with other individuals competent in the occupation. In test construction, the same principles are followed as in other areas of measurement; namely, that of sampling. The larger the number of samples, or items, selected from the various areas of trade knowledge and information, the greater the coverage of the whole content. At the same time, small areas in which a tradesman may not be thoroughly knowledgeable are reduced and will not unduly penalize him in the final test result. When competitors clearly understand why a test that adequately samples the major fields relating to a job is more reliable and more valid than one restricted in coverage, they prefer tests sufficiently long to yield a reasonable appraisal of their competence.

In practice, the length of a test should be determined by the number of items necessary to establish the competence of the individual and to be within physical endurance limits of the candidates. Present practices require on an average from 2½ hours to 3½ hours for the written part. Economic considerations, administrative problems, availability of facilities, and accessibiity on the part of the candidates make it desirable to have the written and performance test conducted on a single day.

The number of test items which make up an adequate sampling varies with each occupation. Any decision as to the number of items are a matter of judgment by both subject content and test experts. Two examples are included in the Appendix serve to demonstrate as to how to determine the actual numbers.

XVI. PERFORMANCE TEST CONSTRUCTION

A. Skills Essential to Competency Evaluation

After the job description, an enumeration of all the factors to be considered in planning and developing a test is in order. Ideally, the tasks chosen for a performance test should be those which will decide success or failure on the job. The test constructor must look carefully at the on-the-job requirements and must resist the temptation of selecting tasks which are relatively simple to observe and score, but which are of lesser importance occupationally. In developing a performance test for a cluster of occupations, care must be taken to avoid duplication of tasks. This can be done effectively by referring to the job analysis and then tabulating the essential skills which every tradesman or technician should be able to demonstrate. The work job should then be designed to test for essential work skills.

B. Process-Centered Work Sample

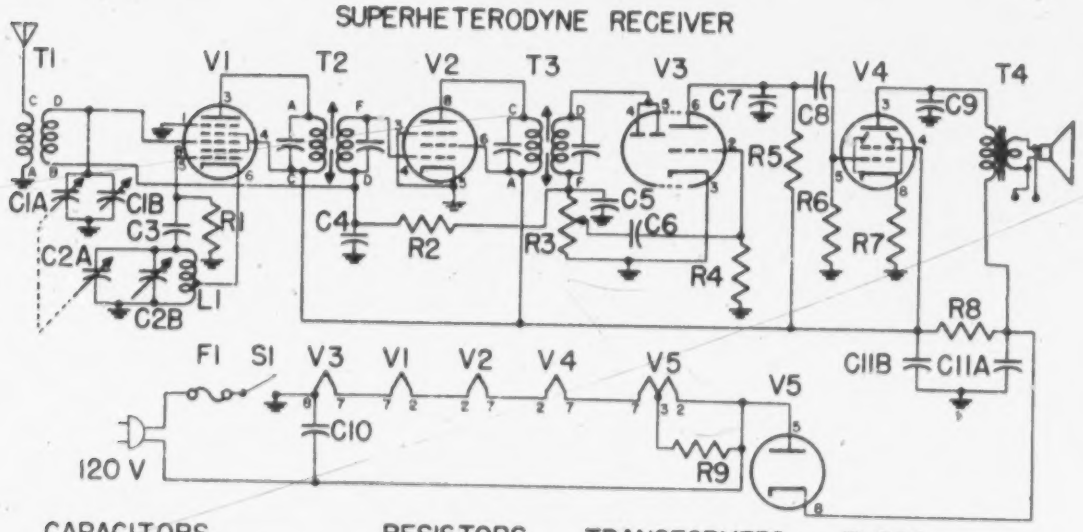
For occupations dealing primarily with procedures (such as the electronic industries--communication maintenance and repair), locating a fault and then correcting it represents the major tasks. The most efficient procedure to locate a difficulty is one that is systematic and logical. A candidate, not thoroughly familiar with all the steps involved in troubleshooting, might eventually find the trouble by chance or by some circuitous or illogical method. The use of systematic, logical and sensible methods supports the assumption tha. the candidate will find troubles in a variety of equipment. In this situation, it is useful and will increase the discrimination and consequent validity of the test if the procedure employed, as well as the method by which the fault is corrected, are evaluated.

Example of a Process-Centered Work Sample:

Figure 2 represents a schematic diagram of an AC-DC five tube receiver. This is a good example of the equipment suitable for a more elementary performance test because it contains concepts and components which are found in on-the-job field operations. While there are, of course, many more complex receivers, the AC-DC receiver is representative to troubleshooting in more complex equipment. It is offered here only as an example of how a task in troubleshooting might be scored and recorded because it fulfills the majority of the occupational task requirements.

The schematic diagram, voltage and resistance chart and block diagram are information available to a repairman on the job.

Figure 2. SCHEMATIC AC-DC RECEIVER



CAPACITORS		RESISTORS		TRANSFORMERS		TUBES	
C1,C2	TUNING CAPACITOR	R1	20 K.OHMS	T1	ANT.	V1	12SA7
C3	.0001 MF	R2	3.3 MEG.	T2	IST I.F.	V2	12SK7
C4	.1 MF	R3	500 K. POT.	T3	2ND I.F.	V3	12SQ7
C5	.0002 MF	R4	3.3 MEG.	T4	OUTPUT	V4	50L6
C6	.002 MF	R5	510 K.OHMS			V5	35Z5
C7	.0002 MF	R6	510 K.OHMS	COIL			
C8	.02 MF	R7	150 OHMS	L1	OSC. COIL		
C9	.02 MF	R8	1500 OHMS				
C10	.1 MF	R9	50 OHMS				
C11A	50 MF						
C11B	30 MF	FUSE					
C12	.2 MF (NOT SHOWN)	F1	3 AMP.				

THIS SYMBOL DENOTES CONNECTIONS TO "B-" BUS

One of the first steps in developing a troubleshooting test is to determine where to start. If the rating includes starting the equipment, checking the fuse, visual inspection of the set, each step offers little or no information for test purpose as to a candidate's proficiency. It was decided, by the test constructor, to have the candidate make only those troubleshooting tests that involved the use of the ohmmeter, voltmeter, and signal generator.

The scoring sheet (Figure 3) was developed after several formal tryouts and is the result of comparing several methods to determine which was practical and discriminating.

Figure 3. SCORING SHEET

P.T.____ CODE____

NAME_____
CLASS____ DATE____
ADM._____

TEST POINT

	R	V	SS
V1			
1			
2			
3			
4			
5			
6			
7			
8			
V2			
1			
2			
3			
4			
5			
6			
7			
8			
V3			
1			
2			
3			
4			
5			
6			
7			
8			
V4			
1			
2			
3			
4			
5			
6			
7			
8			
V5			
1			
2			
3			
4			
5			
6			
7			
8			

COMPONENT CHK.

CAPACITORS	
C1	
C2	
C3	
C4	
C5	
C6	
C7	
C8	
C9	
C10	
C11A	
C11B	

RESISTORS	
R1	
R2	
R3	
R4	
R5	
R6	
R7	
R8	
R9	

TUBES	
V1	
V2	
V3	
V4	
V5	

TRANSFORMERS	
T1	
T2	
T3	
T4	

COIL	
L1	

TIME STARTED____ STOPPED____ COMPONENT: YES_ NO_
STAGE: YES__ NO__ STEPS___ ___ ___

A precise and straightforward check-by-check procedure of relevant troubleshooting steps was derived for each of five troubles. These steps were agreed upon by several experts, each of whom troubleshot the radio receiver independently. Several additional checks (which although not basic were considered to offer some information) were also included.

Each relevant check taken by the candidate was given 20 points; each associated but less relevant check was given 5 points. Time taken to troubleshoot and success in identifying trouble were found to be less discriminating.

It must be stressed that this test and its scoring procedure were developed to meet specific needs. The scoring practices may not be adequate for other testing needs under different conditions.

Every occupation requires its own custom-built tests. The actual assignment for this work-sample procedure testing job was as follows:

> This receiver is properly aligned and in good condition with the exception of a single faulty component. This component must be identified and the trouble corrected. The proctor should be advised when the trouble is located. The evaluation will be based on the method used, the logic employed, the naming of the trouble after a few rudimentary tests, and time required. (Time: _____ minutes)

C. Product-Centered Work Sample

For some occupations like the construction industries or machine trades, the specifications of the job clusters include the application of procedures and processes to the making of a finished part or product. Again, the job analysis is emphasized as a technique for establishing those work skills and areas of information that are critical to the completion of a work-piece or unit which becomes an important measurement of a craftsman's occupational competency.

A detailed analysis for specifying a number of product-centered work samples is illustrated and used in Chapter V for the machine industries occupations.

XVII. TIME LIMITS FOR WRITTEN AND PERFORMANCE TESTS

The nature of the occupation, availability of the necessary facilities, travel time and distance for candidates to reach the testing center are indirect factors which must be considered. Excessive time required may result in fatigue and adversely affect the test results. Present practice in government, the Civil Service, even institutions of higher learning is to limit written examinations to three hours at one time. In unusual cases, they may extend to four hours.

The performance examination can rarely be conducted in less than four hours. More common practice requires between four to seven hours. It may be well to conduct the written and performance tests on two different dates. When a group is large enough, it may be divided with one group engaged in the written examination while the other carries on the performance. Aside from present practice, there are no specific guidelines for establishing time limits. The content and the skills of each occupation and the competency to be demonstrated determine the length of each test.

Chapter V

TEST ADMINISTRATION

I. GENERAL TESTING CONDITIONS

The conditions under which tests of any kind are administered, especially tests of performance, can scarcely be overemphasized. The test room should be free of outside distractions and comfortable as to temperature, air and illumination.

An employee starting a new job and a candidate being introduced to an unfamiliar testing situation, generally, tends to commence work immediately at high speed. Just as newly hired workers are given time to come up to production, it is sound practice to introduce the candidate gradually into a new setting. Simple test tasks to be performed by the candidate and discussed by the examiner, with no credit being given for them serve to ease the tension.

There are good reasons for such practices.

(1) To allay anxiety: People react differently in stress situations and a candidate faced with the prospect of taking a test, which may determine his future, is under some tension. To assist a candidate to relax and enable him to do his best, a warm-up period is desirable.

(2) To acquaint candidates with specific equipment: Small differences in the placement or location of equipment controls, tools, etc. may be confusing under pressure. There is little time during the examination to cover such items.

Instructions must be provided for the candidate which explain the extent and scope of the examination; where he is to report: place, time, date -this should be done by mail. The candidate should receive some written authorization which enables him to enter the examination facilities. He should, also, have adequate time to prepare himself for the examination. In addition, he should be instructed as to what tools, instruments and materials he is expected to bring to the performance test.

The following outlines state clearly what the candidate may expect in both the written and performance examinations for the Machine Industries Occupations (Machine Trades) and Electronic Industries Occupations (Communications).

II. EXAMPLE OF DIRECTIONS FOR MACHINE INDUSTRIES OCCUPATIONS (MACHINE TRADES) CANDIDATES

A. General Information

Occupational proficiency examinations are used, in conjunction with other criteria, for a single or for a combination of purposes:

1. Admission to trade and industrial/technical teacher education programs,
2. To meet state requirements for certification, and/or
3. To establish evidence of occupational competence for advanced standing in undergraduate or graduate programs of study.

They are designed to test the level of skill and knowledge of the candidate in his particular occupational field, as compared with other experienced craftsmen in the same occupation.

Two types of examinations are administered:

The WRITTEN examination, consisting of multiple-choice test items, covers the technical and related information, knowledge and judgements a competent tradesman applies to his work, and

The PERFORMANCE examination includes a variety of typical tasks a competent tradesman is able to perform.

The Scope of each examination lists the areas included in the examination and the proportion of test items from each area.

B. Scope of the Written Examination

Trade information, theory, facts and basic principles the applicant should know:

1. Candidates may take three hours to complete the written examination.
2. Candidates will be notified as to the date, time and place of the examination.
3. A #2 lead pencil must be provided by each candidate.

Percent of the Examination
(12%)

a. Bench and Assembly Work, Layout and Inspection

1) Principles and procedures of layout work
2) Interpretation of blueprints, drawings, sketches
3) Principles of inspection and precision measurement
4) Utilization of precision tools and measuring instruments
5) Specifications from drawings for proper sequence of operations-treatment of materials-determination of tolerances
6) Hand tool, assembly and bench operations and utilization of proper tools

b. Machine Sawing, Filing and Multiple Parts Processing (3%)

1) Principles of operations
2) Special functions - filing - internal and external
3) Cutting tools, blades - friction cutting
4) Selection of blades, file chains, etc.
5) Determination of feeds and speeds

c. Drilling, Tapping, Lapping - (Machines and Attachments) (5%)

1) Sizes and capacities of machines and attachments
2) Single spindle and multiple spindle production operations
3) Principles of precision finishing - reaming - lapping
4) Determination of tolerances and internal finishes
5) Principles and methods of clamping work
6) Determination of speeds and feeds
7) Determination of sizes and kinds of drilling and tapping tools

d) Electrical Discharge Machining (4%)

1) Principles of operation of electronic machining
2) Operation and adjustment of controls
3) Determination of limits and tolerances from blueprints and specifications
4) Accuracy, tooling, size

e) Grinding and Precision Finishing - Processes and Machines (17%)

1) Types and kinds of machines and equipment
2) Principles of operations - form and crush grinding
3) Principles of clamping methods
4) Identification and selection of abrasive wheels
5) Kinds and types of finishes
6) Principles of tool and cutter grinding; surface grinding

f) Materials Treatment, Identification and Testing (5%)

1) Characteristics of materials commonly used in the trades (ferrous - nonferrous materials)
2) Machining characteristics (how to identify)
3) Identification of sizes, specifications and grades of materials
4) Heat treatment methods (hardening, tempering, annealing, carburizing, surface and spot hardening)
5) Special heat treating methods (induction-atmosphere control)
6) Hardness testing methods and instruments
7) Cooling methods - water, oil, air

111

g. Turning Processes and Machines (17%)

1) Types and kinds of machines (engine lathes, turret lathes, automatic turning machines)
2) Speeds and feeds (adjustments and calculations)
3) Clamping devices (mechanical-pneumatic chucks, collets, and face plates)
4) Calculations for tapers, angles, thread turning
5) Accessories - taper attachments - steady and follower rest
6) Tools and instruments for single or multiple operation
7) Methods of internal and external turning - single edge - box tools
8) Production of single or multiple pieces
9) Principles of turret and automatic turning

h. Milling Processes and Machines (17%)

1) Types, kinds and capacities of machines
2) Principles of operation (conventional and climb milling)
3) Accessories and special equipment (vertical heads, dividing-heads, turntables - application and principles of operation)
4) Principles and methods of clamping regular and odd shapes
5) Types, sizes and selection of milling tools
6) Calculation of feeds, speeds for various materials
7) Principles of mechanical or optical contour milling
8) Determination of milling finishes from blueprints and specifications

i. Electrically-Controlled Machine Processes (N/C) (5%)

1) Principle of operation
2) Accessories
3) Steps and procedures in planning the tape
4) Sequences in tool setup
5) Cutting feeds and speeds

j. Shaping and Planing Processes (5%)

1) Principles of operation - shaper - planer
2) Application - advantages and disadvantages
3) Methods of clamping regular and irregular shapes
4) Tools and accessories needed

k. Trade Computations (5%)

1) The use of reference tables, handbooks and charts for calculations
2) The use of formulas characteristic of the occupation
3) Calculations applicable to the work of the occupation

1. Trade Applications of Science (5%)

1) The theory and underlying principles of mechanical advantage
2) Effect of heat and friction
3) Elements of electricity and magnetism
4) Pneumatic and hydraulic power
5) Elements of chemistry, commonly applied to the occupation
6) Other common applications of physical sciences

m. Industrial Illustrations and Drawing Communication (5%)

1) Basic principles of mechanical drawing, shop sketching, and technical illustrating
2) Interpretation of complex drawings, layouts, plans and specifications

C. Performance Examination - General Information

The performance part of the examination provides an opportunity for the candidate to demonstrate his level of proficiency in the manipulative skills and judgements essential to carry on the work of the occupation.

The scope of the performance part of the examination lists the skills, methods and procedures typical of the occupations from which the work sample jobs have been developed.

1. Candidates may take five hours to complete the work jobs.

2. Safety glasses are required and will be provided. Candidates may bring their own glasses.

3. Personal hand tools, such as micrometers and steel rules may be used.

4. Handbooks or reference materials may not be brought into the shop. They will be provided when necessary.

5. Appropriate work clothes should be worn.

6. Adherence to safety standards is mandatory.

D. Scope of the Performance Examination

Trade skills, methods and procedures of operation the applicant should be able to perform:

1. Bench and Assembly, Layout and Inspection Work (15%)

a) Layout of work, including center, reference, contour and dimension lines, surface preparation
b) Using common hand tools and measuring instruments, surface plate, and holding or clamping devices, precision tools and gages

c) Filing, cutting, reaming, threading, fitting, polishing and lapping
d) Testing and inspecting with precision inspection tools -- precision blocks, gages, indicators, hardness testers and comparator

2. Machine Sawing, Filing and Multiple Parts Processing (2%)

a) Set up and weld saw blades
b) Saw to a line
c) Set up and file to a line
d) Clamp regular and irregular shapes
e) Sawing and filing angular and irregular surfaces

3. Drilling, Tapping, Lapping Machines and Attachments (8%)

a) Set up work-vise, straps and bolts, jigs and/or a combination of accessories
b) Drilling, reaming, counterboring, countersinking, tapping and lapping
c) Grind drills

4. Grinding and Precision Finishing (20%)

a) Set up of work - magnetic chuck, vise, sine plate, magnetic accessories and other holding devices
b) Selecting, truing and dressing abrasive wheels
c) Surface grinding
d) Slot or shoulder grinding
e) Form grinding
f) Plain external and internal grinding, taper grinding
g) Cutter grinding - straight and spiral
h) Selection of coolants

5) Turning Processes and Machines (20%)

a) Set up and operate engine and bench lathes
b) Mounting of work, chucks, between center, collets, face plates
c) All basic external and internal turning operations
d) Taper turning, all methods
e) Drilling, reaming, tapping, boring, recessing
f) Turning long and thin stock using follower and steady rest
g) Multiple turning operations
h) Sharpening all cutting tools

6. Milling Processes and Machines (20%)

a) Set up and adjust proper feeds and speeds
b) Set up and operate the dividing head and rotary table; simple and compound indexing
c) Set up vise and other holding devices for holding regular and irregular shapes
d) Set up for various milling operations: slotting, angle milling, spiral milling, graduating
e) Set up for climb milling

7. Electronic Control Operations - Optional (5%)

a) Set up tools in proper sequence
b) Adjust machine for feed and speed
c) Make a tape through numerical operation
d) Mount tape and make a trial run

8. Planing and Shaping (5%)

a) Set up work, tools, and machine accessories for plain surfaces, angular surfaces, contour shapes, slotting and serrating

9. Material Treatment and Testing (5%)

a) Hardening, tempering
b) Annealing, normalizing
c) Packhardening --Carburizing
d) Testing for hardness

III. EXAMPLE OF DIRECTIONS FOR ELECTRONIC INDUSTRIES OCCUPATIONS-COMMUNICATIONS

A. General Information

Occupational proficiency examinations are used, in conjunction with other criteria, for a single or for a combination of purposes:

1. Admission to trade and industrial/industrial/technical teacher education programs,

2. To meet state requirements for certification, and/or

3. To establish evidence of occupational competence for advanced standing in graduate or undergraduate programs of study.

They are designed to test the level of skill and knowledge of the candidate in his particular occupational field, as compared with other experienced people in the same occupation.

Two types of examinations are administered:

The WRITTEN examination, consisting of multiple-choice test items, covers the technical and related information, knowledge and judgements a competent tradesman applies to his work, and

The PERFORMANCE examination includes a variety of typical tasks a competent tradesman is able to perform.

The Scope of each examination lists the areas included in the examination and the proportion of test items from each area.

B. Scope of the Written Examination

1. Candidates may take three hours to complete the written examination.

2. Candidates will be notified as to the date, time and place of the examination.

3. A #2 lead pencil must be provided by each candidate.

1. Basic Electricity as Applied to the Electronic Industry

a. D.C. Circuits (10%)

1) Series

Characteristics and calculations
(voltage drop, total current, power, total resistance, etc.)

2) Parallel

Characteristics and calculations

3) Combinations

a) Series - parallel

(equivalent resistance, voltage drop, total current, power, total resistance, etc.)

b) Voltage divider calculations

b. Measurements, Standards and Tolerances as Applied to Voltmeters, Ammeters, Ohmmeters, and Oscilloscopes (5%)

1) Calculation of shunts for ammeters
2) Calculation of multiplier resistors for voltmeters
3) Meter sensitivity and loading effects
4) Ohmmeter circuitry
5) Oscilloscopes as voltmeters
6) Determination of unknown frequencies with an oscilloscope
7) Observing waveshapes with an oscilloscope

c. Conductors, Insulators, and Semi-Conductors used in Electronic Devices (2%)

1) Factors affecting resistance of copper conductors
2) Factors affecting conductivity of semi-conductor
3) Effectiveness of various materials as insulators

d. Inductive Devices used in Electronics (4%)

1) Transformer operation and calculations
   (phase, voltage, and current relationships, regulation)
2) Radio-frequency inductors
   (rf, if, powdered iron and air core, single and double-tuned)

e. Alternating Current in Electronics (5%)

1) Basic concepts
   (RMS, effective, peak, peak to peak values)
2) Effect of inductance in an AC circuit
3) Effect of capacitance in an AC circuit

f. Single Phase Circuits (13%)

1) Series - R,L.C.
2) Parallel - R,L,C
3) Resonance - Series, Parallel
4) R-L and R-C time constants

g. Batteries as a Source of Voltage for Electronic Devices (1%)

1) Primary cells (zinc-carbon, alkaline)
2) Secondary cells (lead-acid, nickel-cadmium)

2. Electronic Concepts Applied to Vacuum Tube and Solid State Devices

a. Basic Aspects (17%)

1) Static characteristics
2) Dynamic characteristics
3) Biasing
4) Transistor characteristics
5) Other semi-conductor devices

b. Electronic Circuits (31%)

1) Power supplies
2) Vacuum tube amplifiers
3) Semi-conductor amplifiers
4) Oscillators (including multivibrators)

c. Amplifiers in Cascade for RF and AF (8%)

1) Coupling methods
2) Decoupling
3) Frequency response
4) Inverse feedback
5) Gain or attenuation (db)
6) Voltage vs. power amplifiers
7) Optimum transfer of energy

d. Other Concepts in Electronic Application (4%)

1) Amplitude modulation
2) Frequency modulation
3) Pulse definitions
4) Logic concepts

3. Basic Methods and Procedures Applied to Maintenance, Correction and Adjustment of Units, Component Parts (25%)

a. Trouble identification

b. Methods of correction

c. Methods of maintenance and adjustment

Candidates may use a slide rule as an aid in performing mathematical calculations during written or performance examinations. The Allied Data Handbook may be consulted during the performance examination. However, no other notes, reference books or other printed matter may be used.

C. Performance Examination

1. General Information

The performance part of the competency examination provides an opportunity for the candidate to demonstrate his level of proficiency in the manipulative skills and judgements essential to carry on the work of the occupation.

The Scope of the performance part lists the typical maintenance assignments encountered in radio, television, recording and other communications equipment.

a. Candidates may take five hours to complete the jobs.

b. The Allied Data Handbook may be used during the performance examination.

c. Personal tools may be used.

d. Appropriate work clothes should be worn.

e. Adherence to safety standards is mandatory.

f. Speed of diagnosis and efficiency in repair techniques will be part of the candidate's performance evaluation.

2. Scope of the Performance Examination

Candidates must complete six jobs in the areas as listed. Both vacuum tube and solid state equipment will be presented for repair or examination. Problems, which normally occur in service, will be used to test the candidate's ability to make diagnoses and repairs generally employed to correct the faulty circuit.

a. Radio Equipment — Perform two of three jobs

1) Alignment of AM and FM radios
2) Component failures (shorts or opens)
3) Repair of low-power transmitters

b. Television Service — Perform three of four jobs

1) Component failures
2) Adjustment of service controls
3) Sound section repairs and adjustments
4) Sweep circuit repairs and adjustments
5) R-F oscillator tuner adjustments
6) Color receiver field adjustments
7) Trap adjustments (4.5 mh)

c. Recording Equipment — Perform one of two jobs

1) Adjust mechanical malfunctions of record changers or tape transports
2) Knowledge of cleaning or drive mechanisms of record changers and tape transports

Performance of the job tasks will require the use of common hand tools of the trade, including proper alignment tools. The candidate should bring the hand tools to the examination.

The candidate must be familiar with the use of the necessary test equipment to perform the assigned work, such as: the VTVM, VOM, oscilloscope, various radio and audio frequency generators, signal tracers, and dummy loads. Problems involving tube failures will not be included in any one of the test items. The knowledge of the use of specific equipment for color television setup may be required. Equipment will be furnished.

IV. INSTRUCTIONS AND DIRECTIONS FOR EXAMINERS

A. Examiners should be furnished with a list of names and Social Security Numbers of candidates and the specific tests to be given.

B. Candidates should identify themselves by their letter of notification.

C. A candidate who fails to present his identification may be admitted to the examination, if his name is on the list.

D. Candidates who appear more than 45 minutes late may not be admitted. (Under exceptional conditions, admission to the examination may be granted and the admission time extended.

A. Directions for Administering Written Examinations

1. Things for the Proctor to do Prior to the Examination

a. Obtain specific directions relative to proctoring the written examination.

b. Obtain list of candidates expected for the examination.

c. Obtain examination papers, answer papers, and other material needed for the test.

2. Before the Examination Begins

a. Check condition of the examination room for proper heat, light, ventilation and adequate seating facilities, and enter any irregularities on the Examination Report.

b. Require candidates to identify themselves as they arrive.

c. Seat candidates taking the same examination at a distance from one another.

d. Candidate may bring only those supplies specifically required for the examination.

e. Check to be sure each candidate has a pencil.

f. If other supplies, such as Handbooks or slide rules are suggested, the proctor should have a reserve supply available.

3. Distribution of Examinations and Oral Instructions

a. Be sure all candidates are seated and ready to begin work.

b. Make any special announcements that may be required.

c. Hand to each candidate the appropriate examination envelope, requiring the envelopes be kept closed until the "start" signal has been given.

d. Say to the candidates, "Take the examination out of the envelope. (Pause) Put your Social Security Number on the line for the Name or Number and today's date, which is _____. Do not put your name on the booklet, unless you do not know your Social Security Number. (Pause) Make sure your Social Security Number is entered on all work you hand in to the examiner.

After you read the directions, begin work. You have three (3) hours during which to complete the examination. Mark your answers carefully in the spaces provided on the answer sheet. If you make an error, erase your mark completely. Failure to do so, may result in an incorrect answer being recorded. Also, note that the answers are in letter form and are arranged from left to right (or in whatever form they may be arranged). After completing the examination and are ready to check out, place all your work and papers into the examination envelope. You may not leave the room for any reason, without first checking with me (or the proctor)."

4. What to do During the Examination Period

a. Note the time when the examination begins.

b. Only one candidate may be permitted to leave the room at any one time.

c. No candidate should leave the room, after another person taking the same test has checked out.

d. Be courteous if questions arise, but do not interpret the intent, or meaning, of a test question, or furnish any information to the candidate. Suggest that the candidate answer the test item as best he can.

e. If questions arise concerning possible errors in items (e.g. spelling, wording, ambiguity or errors in the alternatives), do not correct the error or announce it to the group. Again, tell the candidate only, "I cannot give you any information. Answer the item as best you can." (Do this even if there is no item error.) However, enter any possible error on the Examination Report in the Irregularities Section.

f. Be alert and circulate quietly around the room, continuously remaining in the room to supervise the examination.

g. Stop work at the expiration of the allotted time, except for those who have official permission to continue.

h. As the examinees check out of the examination room, be sure that

(a) the Social Security Number is on the examination booklet and all other papers, and

(b) all test materials are put into the candidate's envelope.

i. Return all examination papers to the appropriate authority.

B. Directions for Administering Performance Examinations

Performance examinations require far greater preparation than written examinations. It is absolutely essential that the following steps be carefully followed. Deviations from these procedures will result in difficulties in administering and scoring the examination. As you complete each of the following steps, check the item off on the Examination Report form (page 109) and, also, record any problems or irregularities. In this way, you will be adequately prepared for the examination, and will avoid needless hardships and time-consuming delays.

If you encounter any difficulties which you cannot handle in preparing for the examination, contact the appropriate person immediately, so that the problem can be solved prior to the test date. _____ _____ will check, prior to the date of the examination, to be sure that everything is ready and will, also, check with you during the actual administration of the performance examination since unexpected problems may arise.

1. Facilities for Examinations

a. Performance tests should be conducted in shops or laboratories equipped with appropriate and sufficient machines, equipment, hand tools and materials. Shops may be located in vocational schools or selected industrial establishments.

b. Specific directions for conducting examinations for each occupation will be placed in the hands of the examiner well in advance of the date of the examination.

c. In any case, the examiner should analyze the test problems to determine amount and type of tools, equipment and other items needed for the number of persons scheduled to take the test. When the best available facilities are inadequate, arrangements should be made to supplement those at hand by securing additional equipment in order to conduct the test under the best possible conditions.

2. Supplies for Testing

Prior to the examination, the examiner should determine the materials and supplies needed to conduct the tests. In cases where supplies are not readily available and must be purchased, the examiner should consult _____.

3. What to do Before the Day of the Examination

a. Read the "Directions to the Examiner" and "Performance Examination" for the examination being given.

b. Check to verify arrangements for time and place of the examination.

c. Visit the testing center to see that all equipment, tools, etc., necessary for the test is available.

d. Check all equipment to make sure that it is in good working condition.

e. Set up all tools, machines, equipment and working conditions so each working station is as nearly identical as possible.

f. Prepare all partially completed test pieces and/or defective equipment for troubleshooting for the examination.

g. Check to make sure that all tools, materials and supplies are in a convenient location for each examinee.

h. Place all test materials (booklets, answer sheets, scrap paper, rating forms) in a large envelope for each candidate, and mark it with the candidate's name and Social Security Number.

i. Familiarize yourself with the scoring procedure.

THESE STEPS ARE TO BE COMPLETED AND RECORDED ON THE "EXAMINATION REPORT: PRELIMINARY CHECK-LIST FOR PERFORMANCE EXAMINATIONS: BEFORE THE DAY OF THE EXAMINATION." (See sample form page 107)

4. What to do During the Examination

a. Before actual work begins, permit the examinee to acquaint himself with his surroundings since the shop may be strange. Point out the location of such facilities as: electrical outlets, service facilities, and other physical details.

b. Note the time each job is started and finished by each candidate.

c. Observe the candidate at his work. Check his work methods quietly and inconspicuously. Keep a record for your own purposes as the examination progresses.

d. Spend approximately the same amount of time with each candidate.

e. Check and rate each step of the job as it is completed by each candidate.

f. Tag or stamp each completed job with the candidate's examination number.

g. Maintain an impartial attitude at all times. It eliminates criticism and aids in reducing tension which, quite naturally, develops during the examination.

h. Provide Handbooks, reference tables, etc. required in the examination.

i. Carefully record all ratings required on the "Performance Rating and Summary" form for the specific field. If there is no special form, use the General Performance Rating and Summary form. (See Figure 4, page 107.)

5. What to Avoid During the Examination

a. Allow only one examinee to leave the room at any time.

b. No work is to be taken from the shop during the examination.

c. Avoid interpreting the meaning or intent of a job assignment. Suggest that the candidate proceed according to his understanding.

d. Rating sheets should be marked only when out of the candidate's presence.

e. Refrain from giving any information to a candidate concerning his progress or the probable result of the examination.

f. Such matters as a candidate's occupational experience, his education, or personal problems should not be discussed.

6. At the Close of the Examination

   a. Stop all work at the expiration of the maximum allotted time.

   b. Collect all examination forms and directions.

   c. Identify all test pieces with the respective candidate's number.

   d. Retain all finished work.

7. Completing the Scoring of the Examination

   a. Most of the performance ratings can be made during the examination.

   b. Inspect and rate all other completed work immediately after candidates have departed from the shop.

   c. Use only the performance rating sheets provided.

   d. Enter comments which have a bearing on the competency of the candidate on the back of the sheet. This is particularly important in cases where the performance is below average.

   e. Sign the rating sheet in place provided for "Examiner".

   f. Return all scoring sheets, question papers, prints and other materials provided for each examination.

   g. All totals will be computed at the National Occupational Competency Testing Project Center.

Figure 4. SAMPLE EXAMINATION REPORT

PRELIMINARY CHECK-LIST FOR PERFORMANCE EXAMINATION

Complete each step as listed, filling in the date when completed.

1. _____ Have read "Directions to the Examiner" which are specific to the examination being given.

2. _____ Have read the "Performance Examination", which is the actual examination the candidates take.

3. _____ Arrangements for the time and place of the examination have been verified with the Regional Coordinator.

4. _____ Visited testing center to be sure that all equipment; materials, and tools necessary for the test are available. (See: "Directions to the Examiner" and the "Performance Examination" for the specific items that will be needed.)

5. _____ All equipment has been checked and is in good working condition.

6. _____ Tools, machines, equipment and working conditions have been set up, as nearly identical as possible, at each work station.

7. _____ All partially completed test pieces and/or defective equipment for troubleshooting are prepared and ready for the examination.

8. _____ Tools, materials, Handbooks and supplies are in a convenient location for each examinee. (This should be similar to the typical occupational work situation where he would select equipment from a choice of tools and supplies normally available. However, the examinee should not be required to spend time in securing stock and performing other preliminary work.)

9. _____ Job assignments, necessary drawings and other written directions, relating to the test, have been placed in an envelope and marked with the candidate's name and Social Security Number. Partially completed items are available for the candidate.

I have completed the above steps, completely familiarized myself with all procedures I am to follow during the examination and have verified this with the Regional Coordinator.

Signature _____ _____
(Date)

(If there are any problems relating to these steps, or if you see any problems in administering the examination, contact the coordinator of the testing program.)

Figure 5. SAMPLE REPORT FORM OF IRREGULARITIES

OCCUPATIONAL PROFICIENCY EXAMINATION: _____
(Title)

Center _____ Date _____

Name	Social Security Number	Attendance		Individual Irregularities*
		Present	Absent	

Time examination began: _____ Time examination completed: _____

Possible Item Error: (Remember: Candidates are not to be told to change the item!)

* Irregularities affecting all candidates: (e.g. overcrowding, power failure, distracting noises, emergencies over which you have no control, etc.)

_____
_____
_____

Examiner's Signature _____ _____

Chapter VI

WRITTEN TEST RATINGS AND PERFORMANCE EVALUATION PROCEDURES

The work of the tradesman involves the use of instruments for precision measurement. Tests are, likewise, measuring devices to provide accurate information regarding an individual's occupational competence. To achieve accuracy, the manner in which test results are recorded requires careful treatment.

I. RECORDING TEST ANSWERS - WRITTEN TESTS

Modern data processing equipment makes it possible to record, tabulate, and rate test results with remarkable speed. Processing with this equipment, also, sets a pattern for recording the answers and solutions to test items of written examinations. There is no one precise form for all examinations because of the many diverse methods of recording which are used. Whatever method is employed, precise directions must be given to the examiner on how to proceed. For the development of a recording form, the following must be taken into consideration:

A. <u>For Machine Cards</u>

1. What are the limitations of the equipment?

2. How many spaces are available on one card when scanning equipment is used?

3. What physical arrangement is required for the layout of the card?

B. <u>How Should Test Results Be Reported?</u>

1. What kind of computer facility is available?

2. What kind of program must be planned to provide:

   a) Listing of candidate's data (desired background)
   b) Grades in numerical or letter form
   c) Statistical treatment of data:

      1) Item analysis
      2) Difficulty level and discrimination analysis
      3) Evaluation of various parts of the examination
      4) Determination of correlation coefficients
      5) Determination of control tendencies and standard deviation
      6) Other information desired about the test or the relative standing of the individual

Data processing equipment and facilities are continuously improved. Before a method of recording is established, it would be wise to examine existing procedures in the field for the latest equipment available. Recent developments permit the use of larger answer sheets - 8½ x 11" instead of the smaller IBM card. The guiding consideration should be to strive towards objective rating by objective, mechanical means.

II. RECORDING TEST RESULTS - PERFORMANCE TESTS

Evaluating the work performance and recording the results of a performance test is more difficult because subjective judgements must be recorded as part of the overall evaluation. Determining factors for developing a rating form are the nature of the trade, the kinds of skills and judgements to be evaluated, the relationship between objective test items and subjective observations, the ease or difficulty in observing and rating the various aspects of the performance, etc.

Figure 6. TROUBLESHOOTING EVALUATION FORM

Place _____ Name _____

Examiner _____ Test Point Date _____

Step-by-Step Sequence Check

VI.	R	V	SS
1			
to			
:			
8			
V2			
1			
2			
to			
8			
V3			
1			
2			
to			
8			
V4			
1			
2			
to			
8			
V5			
1			
2			
to			
8			
Total			

Component Check

CAPACITORS		
C1		
to		
:		
C11B		
REGISTERS		
R1		
to		
:		
R9		
TUBES		
V1		
to		
:		
V5		
TRANSFORMERS		
COIL		
L1		

1) A precise check-by-check procedure of relevant troubleshooting steps were derived for each of five troubles (not necessarily including 8 items).

2) Each relevant step was assigned 20 points.

3) Each associated step, although not basic was assigned 5 points.

4) All irrelevant steps were assigned 5 points (to be subtracted from the total).

A. Evaluation of Troubleshooting Items

If troubleshooting represents the major portion of the test, a listing of all the steps involved should be developed, including a space for check off. Possible values for each step completed or not completed should be included. A separate rating scale may be set up for the performance of the connected unit. A step-by-step evaluation form follows.

B. Evaluations Including Special Time Factors

Should speed be a vital factor, then the maximum time permitted should be carefully determined and penalties indicated for those candidates who exceed the time limits. For candidates completing the tasks in substantially less time premiums for early completion must be indicated. A rating form for evaluation of speed follows:

Figure 7. TIME STANDARDS FOR JOB ____

Place _____ Date _____

Examiner _____ Name _____

Time/Minutes	Rating	
25	90	
30	85	
35	80	
40	75 standard time	
45	70	
50	65	
55	60	
TOTAL SCORE		

C. Evaluations Including Measurements of Accuracy

For instance, if a part must be machined to a tolerance of $\pm$ 0.001" on a dimension of, say, 7.000", the candidate who machines the part between 6.999" and 7.001" may receive a perfect score. A candidate who exceeds these maximum tolerances would score below a passing grade. Other ratings between 100 and 70 should be established in advance for measuring dimensional accuracy to identify scores to be assigned for gradation in accuracy.

If the same examination stated that one objective was to rate the candidate's competence to machine to precision limits of one ten thousandth part of an inch (0.0001") and closeness to the base dimension was

important, then a different rating scale might be applied. The candidate who machined the part exactly 7.000" would receive a perfect score. The ±.001" tolerance might be divided into ten increments representing a perfect score of 100 and a passing score of 70. Thus, the candidate who machined the part to 6.9999" (-0.0001") or 7.0001" (+0.0001) would receive 100-3 or 97; 6.9998 or 7.0002", 94, etc.

Also, for those exceeding the limits of tolerance (± 0.001") a series of scores could be established below the passing score to indicate limitations for accuracy of machining. Thus, the scores permit differentiation between those with occupational competence to work to the nearest ± 0.0001" and others whose skill is limited to ± 0.001".

The same techniques may be applied to other trade or technical occupations which may not require the same degree of precision measurement. Directly measurable items have the advantage of objectivity and fairness. Such items as dimensions, location and sizes of cut-away sections, sizes of threads, etc. can be measured. The important point is that tolerances should be given and values assigned to represent a range of ratings which differentiate the capable occupationally competent craftsman from a less skilled worker.

Figure 8. DIMENSION RATING FORM

Place _____ Job _____ Date _____

Examiner _____ Name _____

Dimension	Specimen Dimension	Weight	Rating	Remarks
A		5		
B		10		
C		10		
D	(as needed for each work job)	10		(as needed for each work job)
1				
1				
1				
K		15		
Threads		15		
Workmanship		20		
Usability		15		
	TOTAL	100		

D. Evaluations of Finished Product

For finished product emphasis, careful distinction must be drawn between methodology employed and the items of the finished part to be measured. Every effort must be made to define such terms as workmanship, work habits, work methods. Sometimes, a large number of factors are assumed under the heading "skill". The latter may cover analysis, diagnosis and planning the job, selection of appropriate materials, tools, instruments, etc. Often, the category also includes correctness of procedures, interpretation of drawings and other directions. The list could be extended considerably.

Obviously, an evaluation expressed in a single number on a rating scale, or an identification as; unacceptable product, marginal quality, average, good, quality product would be extremely subjective and would inevitably lead to substantial differences among the examiners.

E. Evaluating Proper Work Methods

The completion of the product indicates certain occupational competence. In addition (for tradesmen interested in teaching), work habits acquired over a period of time and the methods employed in producing the product are significant parts of overall competence. Scoring forms to evaluate procedures and work habits might be developed in several ways.

1. A common type of rating form lists each step required for a particular job. The form illustrated in Figure 9 simply requires a "Yes" or "No" check that later is considered as indicating whether or not the candidate follows correct work steps or procedures.

Figure 9. RATING FORM FOR CHECKING SEQUENCE OF STEPS

Trade _____ Date _____

Examiner _____ _____
Candidate Identification

	YES	NO
Step 1 _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _	___	___
Step 2 _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _	___	___
Step 3 _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _	___	___
Step 4 _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _	___	___
Step 5 _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _	___	___
Step 6 _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _	___	___

2. A different evaluation form may list a series of statements of the steps employed, followed by columns to record a number or letter grade to indicate how well performed and safety precautions followed. The examiner evaluates the performance against each statement and records his evaluation. The overall judgements are then totalled.

3. Still another form may be used to evaluate particular phases of a job. While a job may be completed within all tolerances, the finished product may not show how many additional steps or operations had to be performed before the final cut was taken to get to the required dimension. In such a case, each additional cut or operation is noted and points are subtracted for this extra work.

4. Sometimes, a simple graphic scale, with values from poor to excellent, merely requires the examiner to note the term which best describes the candidate's performance on an operation or series of work processes.

Example: Clamping Work Piece

	1	2	3	4	5
Level of Performance	Unfamiliar with Process	Unsure of Steps	Average Performance	Capable: Above Average	Excellent Craftsman Techniques

Such items as safe practices, orderly arrangement of tools and materials, machine set ups, and many others can be arranged in similar fashion.

Attempts have been made to design one rating form for all trade performance tests. However, desirable this may seem to be administratively, the differences among the various trades, their special charactistics, and the need for objectivity requires a special rating form for each occupation. The two forms developed for the National Occupational Competency Testing Projects for the Machine Industries Occupations (Machine Trades) and the Electronics Industries Occupations (Communications) appear is Figure 10 and Figure 11, respectively. Note that on Figure 11, all jobs are recorded on one rating form. The final score will be the sum total for all rated jobs completed divided by the number of jobs. A suggested rating distribution is provided on the back of the form. The examiner is also asked to express his judgement, based on observations, of the candidate's performance.

F. Examiner Reliability

Despite all efforts to eliminate subjective judgements of examiners in evaluating the performance of a candidate, certain job factors cannot be appraised in any other way except by personal judgement. Experience has shown that training examiners with proper focus on the items to be evaluated and rating scales can produce agreement among different examiners. A high factor of reliability among examiners evaluating performance tests has been found when these conditions are met.

Figure 10 SAMPLE EXAMINER'S PERFORMANCE RATING SCALE

NATIONAL OCCUPATIONAL COMPETENCY TESTING PROJECT
OCCUPATIONAL COMPETENCY EXAMINATION
MACHINE INDUSTRIES OCCUPATIONS-MACHINE TRADES

Candidate's Name _____ _____

Location _____ Examination Number
Date _____

While the candidate is performing the assigned tasks, rate his performance according to the scale indicated for each subdivision. The top rating would be equivalent to an extremely able or competent worker. A good or above average worker would be rated 2 or 3 points below the top rating. The average worker rates at the midpoint. Workers with below average performance would rate 2 or 3 points below the midpoint, and the inept worker rates to zero. After he has completed the task, grade his finished product according to the dimension specified: deduct points for variations beyond normal tolerances. Check any weaknesses in box provided.

Bench and Layout Inspection

Observed Performance	Rating 0-10	Finished Product	Rating 0-10
() Handling of layout tools		() Accuracy	
() Planning of layout procedure		() Dimension lines	
() Preparation of surfaces		() Centers of holes	
() Filing - center marking		() Sizing holes	
() Tapping			
() Other tools	( )		( )

Turning Processes and Lathe Work

Observed Performance	Rating 0-10	Finished Product	Rating 0-15
() Set up of work		() Accuracy	
() Set up of tools		() Finish	
() Set up of accessories		() Knurls	
() Operation of lathe		() Other	
() Feed - speed	( )	() Tool bits	( )

Milling Processes and Machines

Observed Performance	Rating 0-15	Finished Product	Rating 0-15
() Set up of work		() Accuracy	
() Set up of cutters		() Finish	
() Set up of accessories		() Squareness	
() Feed - speed		() Appearance	
() Safety procedures	( )	() Burrs, marks, etc.	( )

Heat Treatment

Observed Performance	Rating 0-5	Finished Product	Rating 0-5
() Proper handling of part		() Proper hardness	
() Quenching procedure		() Proper tempering color	
() Tempering procedure			
() Safety	( )		( )

Drilling and Drill Press Work

Observed Performance	Rating 0-5	Finished Product	Rating 0-5
() Manner of clamping work		() Size of holes	
() Use of vise-parallels		() Location of holes	
() Use of tapping attachment		() Removal of burrs	
	( )	() Drills	( )

Grinding Processes and Machines

Observed Performance	Rating 0-15	Finished Product	Rating 0-15
() Set up of work		() Accuracy	
() Truing of wheel		() Finish	
() Dressing of wheels		() Squareness	
() Sequence of steps		() Formsand radii	
() Selection of wheels			
() Feeds and speeds			
() Safety practices			
() Coolants	( )		( )

Planer - Shaper Processes and Machines

Observed Performance	Rating 0-8	Finished Product	Rating 0-8
() Set up of work		() Accuracy	
() Set up of tools		() Finish	
() Feeds and speeds		() Other	
() Safety factors	( )		( )

Rating - Observed Performance _____

Rating - Finished Product _____

Rating - Overall _____

Examiner's Signature _____

(Give your personal impression of the candidate on the back of the sheet)

Figure 11

NATIONAL OCCUPATIONAL COMPETENCY TESTING PROJECT
OCCUPATIONAL COMPETENCY TEST
ELECTRONICS INDUSTRIES OCCUPATIONS (COMMUNICATION)

EXAMINER'S PERFORMANCE EVALUATION SCALE

Candidate's Name and Number _____

Place _____

While the candidate is performing the assigned tasks, rate his performance on each job on the scale shown below. Indicate any weaknesses in the left column.

The candidate should be rated on two features of his performance: a) his work methods, and b) the finished job. The following seven aspects of his work methods are to be used:

1. Does he approach the identification and location of the trouble systematically in a step-by-step manner?
2. Does he use the accepted trade method in correcting the trouble?
3. Does he select test equipment properly?
4. Does he perform hand-tool operations skillfully?
5. Does he work neatly, accurately and safely?
6. Does he work at an acceptable speed?
7. Does he avoid excessive and inappropriate use of reference material?

In assigning ratings to the job performance, remember that the scores indicated are maximum and should be given only to the extremely competent candidate (for a breakdown see next page).

Work Methods and Finished Job	Maximum Score	Jobs								Comments: Weaknesses
		1	2	3	4	5	6	7	8	
a-1 Procedure in identifying and locating trouble	20									
2 Method of correcting trouble	15									
3 Selection and use of test equipment	10									
4 Skill of hand tool operations	10									
5 Neatness, accuracy and safety of work	5									
6 Speed	5									
7 Use of reference materials	5									
SCORE FOR WORK METHODS	70									
b- SCORE FOR FINISHED JOB	30									
Total Score	100									

Final Score = $\frac{\text{Sum of the scores on all jobs}}{\text{Number of Jobs}}$

Final Score: _____ Examiner's Signature _____

These are estimated average values. A rater should go above or below these figures to more accurately record his judgement.

Figure 12. BREAKDOWN OF FIGURE 11

Level of Competence	Rating According to Value of Maximum Score				
	30	20	15	10	5
Worker inept or out of practice	6	4	3	2	1
Worker with below average and limited facility	12	8	6	4	2
Average worker in the field	18	12	9	6	3
Good or above average worker	24	16	12	8	4
Extremely able, highly competent worker	30	20	15	10	5

State your overall impression of the candidate's competence as a worker in the field, in addition to the numerical score.

Chapter VII

EVALUATING OCCUPATIONAL COMPETENCY TESTS

When a craftsman has to take measurements requiring close tolerances, he uses a precision measuring instrument, gage, etc.. Repeated measurements of the same dimension or unit produce the same result. Therefore, it is a reliable instrument. Electrical, mechanical, hydraulic, and other measuring devices function continuously within certain limits of accuracy and are reliable.

On the other hand, measuring occupational competence involves many more variable factors, is far more complex, and the criteria for which competence must be measured objectively are much more difficult to establish. However, methods and techniques have been developed which assist the test constructor to check the quality of the measuring instrument.

A test must measure not only what it is intended to measure, but it must also be valid for the purpose it is to serve. Occupational competency tests are intended to measure the proficiency level of an individual relative to the proficiency standards demanded in the occupation.

In some states, the number of years of experience an individual has spent in his field and an evaluation of job performance by supervisors and employers is used as an indication of job competence. The subjective nature of this approach has led to a more positive direction for determining criteria against which occupational competency is to be measured. Specific information for test development is identified from job studies, job descriptions and comprehensive analyses of occupational job clusters. Occupational skills and essential information which are characteristic of the proficiency at a certain level of performance have been identified and serve as the criteria of what a competency test must measure.

This chapter deals with the evaluation of occupational competency tests in terms of validity, reliability, analysis of test scores, difficulty level and discrimination of test items, length of tests, item analysis, and test standards.

I.. KINDS OF VALIDITY

Tests are used for a variety of purposes, each requiring a certain type of validating evidence. For purposes of this report, four types of validity will be discussed. These represent accepted practices by the American Psychological Association, the American Educational Research Association, and the National Council on Measurements Used in Education.

A. Content Validity

Content validity is most appropriate with achievement evaluation. The question to be answered is primarily, "Does the test measure the skills and technical information which are characteristic of the competent craftsman?". A test has content validity when it contains a representa-

tive sampling of the manipulative skills, the trade information, and "know-how", at the journeyman level. Content validity relates to the adequacy of the test, as being representative of an area of behavior, such as: performing certain tasks within time and accuracy limits. Content validity is a matter of the occupational specialist who has the overall knowledge and skill of the occupation making a determination on content adequacy.

B. Concurrent Validity

Concurrent validity is a relevant concern in the evaluation of achievement. It represents a comparison of test based categories and actual achievement data in the occupation. Concurrent validity is, thus, expressed in terms of the relationship between test performance and an accepted contemporary criterion. For occupational competency tests, this represents a very difficult problem. At this time, no national performance criteria or standards are available. Instead, the computed correlation coefficient, based on test results, provides a quantitative measure of this relationship.

C. Predictive Validity

The predictive validity of an instrument, requires criterion against which predictions can be compared. While much work has been done in the academic subject content area, and tests for supposedly predicting college success are widely used, occupational competency tests are not intended to measure future success as teachers. They serve only to establish an individual's proficiency in his occupation.

D. Construct Validity

Construct validity concerns the extent to which a test indicates the meaningful characteristics of the occupation. Information about such characteristics provides an understanding of the tradesman's performance level on the job. The content validity of an occupational competency test is judged on the basis of how adequately the test samples (represents) the total content of the occupation. Construct validity, however, concerns the test's ability to measure the individual's actual achievement in the occupation. A competency test of high construct validity must distinguish between a tradesman and others performing at different levels of capability.

The following kinds of evidence are useful in establishing the construct validity of a measuring instrument:

1. Expert judgement concerning the extent to which responses to the test items provide knowledge about the individual's possession of the skill and information;

2. Correlation with other tests, particularly those that are accepted measures of the same construct;

3. Correlation with other characteristics of the individual; and

4. Correlation with factors in the individual's environment which might affect test performance (types of employment, kinds of equipment different from test facilities, etc.).

In essence, the focus of construct validity is upon the validity of the specifications derived from the occupational analysis. For instance, if occupational competency is to be measured for the person who is to teach in the machine trades, such characteristics as operating skills and trade knowledge related to the various machines, tools, apparatus, and materials, methods of work, etc. (which are characteristic of a highly competent individual), must be included in the test.

II. VALIDATION OF TESTS

A. Logical Approach

In occupational competency testing, two questions are of major concern. The first, "Is the test suitable?" and the second, "Will the test scores be reliable?" Discussing the various aspects of validity and reliability is one thing; proving that the test is valid is much more involved.

Mere scores, as such, have little meaning. A candidate may achieve a score of 80, 60 or 90. Does this mean he is 80%, 60% or 90% proficient? Actually, each score represents merely a record of achievement on a sampling of acceptable occupational competence. Any judgement regarding competence is an inference from this score which reflects only the number of correct answers. Its validity is not self-evident, but is something that must be established on the basis of adequate evidence. This evidence is derived from 1) what an occupationally capable individual needs to know and 2) what he must be able to do in the occupation, according to certain standards of performance and craftsmanship. Critical examination and judgement then form the basis for measuring knowledge and performance. Statistical procedures for measuring content validity are not now in existence. Content validity rests nearly always on some form of expert judgement.

B. Emperical Approach

In contrast to the logical approach, the emperical method provides different kinds of information. The following simple example provides identification of whether each item on an occupational competency examination does or does not measure individually what is expected of the test as a whole. By listing the test items in numerical order and calculating the percent of the total number of candidates who answer each one correctly, those items which discriminated and those which failed to do so may easily be established (Table X).

Table X. ITEM ANALYSIS OF RESULTS OF A CARPENTRY EXAMINATION BY DIFFERENT CRAFTSMEN

Item	Percentage of Selected Groups Passing the Item	
	Carpenter	Other Construction Workers
1	90%	85%
2	60%	65%
3	85%	26%
4	25%	48%
5	76%	30%
-	--	--
-	--	--
-	--	--
200	25%	28%

In this case, the written test for carpentry, containing 200 items, was administered to a group of journeymen carpenters and journeymen in other construction occupations. From these scores, it becomes evident that items 3 and 5 are unique to the carpentry trade. Items 1, 2, 4 and 200 do not discriminate between carpenters and other construction workers. From an empirical approach, using an item-by-item analysis, the value of each test item to measure specific occupational competencies may be established.

III. ANALYZING PERFORMANCE TEST SCORES

A. Steps in Making the Analysis

1. The first step in analyzing a performance test score is to put the test scores into a distribution. This is done to derive more meaning from them. Actually, a test should be evaluated by asking three questions:

   a) Did the test discriminate? Did all the testees achieve about the same score or was there a variety of scores on the test? Do the test results show clearly as to who was the best performer, the next best, and so forth?

   b) How difficult was the test? The frequency distribution will reveal that, if nearly all of the group achieves high scores, the test was too easy. When most examinees obtain very low scores, the test might well be too difficult.

   When the examinees' scores cover a wide range, with some individuals getting high scores, some low scores, but the majority scoring in the middle range, the test reflects a proper difficulty level.

c) Did the test accomplish its mission? Whether a test has fulfilled its purpose can be quickly determined from a score distribution chart (Table XI).

Table XI FREQUENCY DISTRIBUTION OF SCORES

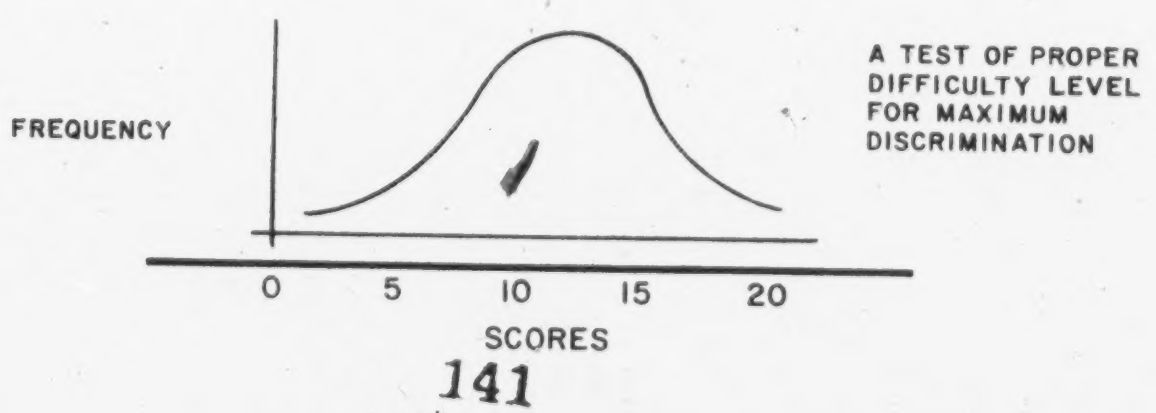
Raw Score	Tallies	Frequency
8	1	1
7	11	2
6	111	3
5	1111	4
4	11	2
3	11	2
2	1	1
	Total Number: 15	

The scores are listed from the highest to the lowest in a vertical column. The number of examinees who achieved each score are tallied in the middle column. One examinee has achieved a score of 8, four have scores of 5, and one a score of 2. Obviously, this test discriminates among those who took the test.

B. Difficulty Level of the Test

For a visual evaluation of the frequency distribution it is helpful to show the distribution graphically as a curve. This may be done by plotting the raw scores from the test along the horizontal axis, and the number of individuals taking the test along the vertical axis. (Figure 13)

Figure 13. CURVE OF TEST OF PROPER DIFFICULTY LEVEL FOR MAXIMUM DISCRIMINATION

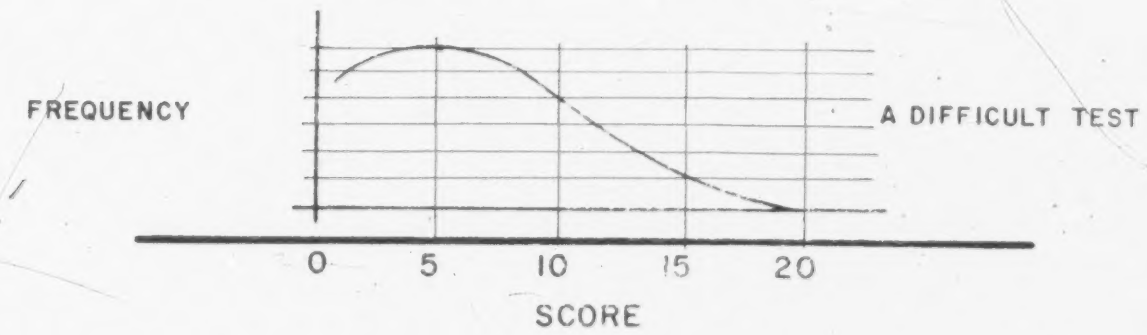


141

The bunching of scores near the middle of the scale simply shows that this is where most of the examinees scored. There are neither a lot of individuals showing high competency nor a disproportionate number of less competent individuals. Generally, this kind of curve describes a test that is at the proper difficulty level for the individuals being tested.

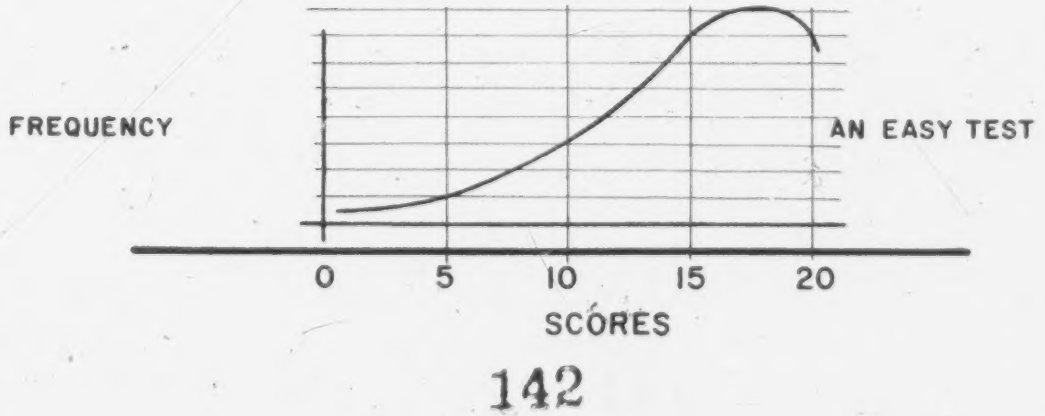
Referring to the graph in Figure 14, it is apparent that the examinees did not do very well in the test since most achieved low scores. If the test was well constructed, it may mean that the test was too difficult or the examinees were not qualified. However, as the graph is interpreted, the test was difficult for those who took it.

Figure 14 CURVE OF A DIFFICULT TEST



The opposite condition seems apparent from Figure 15. Very few low scores were achieved as the majority of the examinees received nearly perfect scores. This test was easy for the individuals tested. In order to discriminate better, the test items and test should be made more difficult.

Figure 15 CURVE OF AN EASY TEST



142

The fundamental principle of testing is that there are differences in individual proficiencies. The tests are to be constructed to obtain scores that indicate these differences.

While these procedures provide a relatively simple and quick method of test result evaluation, there are other (statistically more involved) methods which are later reviewed.

IV. DETERMINING TEST RELIABILITY

A. Factors that Determine Reliability

Although validity is one of the major considerations in assessing the quality of a test, its reliability is equally important. A test which is valid must, also, be reliable. Reliability is the consistency with which the test yields the same results in measuring what it does measure. However, a test may be reliable and not valid. To exaggerate, for instance, if height is used to measure intelligence, it will be reliable because height measured with a yardstick is very constant but will not be valid for the purpose of measuring intelligence. Reliability pertains to a class of test characteristics. These are stability, equivalence and internal consistency.

1. Test and Retest Method

In the test and retest method, the test is administered to a group of individuals representative of the group to which the test will eventually be administered. After a suitable lapse of time, say four to six weeks, the same test is administered to the same group and the correlation coefficient between the first and second administration of the test computed. This coefficient then serves as a reliability estimate. It is called "coefficient of stability".

While there are some advantages in the test and retest method, there are also some limiting factors. The choice of appropriate time interval is a critical factor. Additional limiting factors, such as: availability of a suitable test population, time of the examinees, accessibility to testing facilities, etc., make this method of limited practical value.

2. Administration of Equivalent Forms

The equivalent form of an occupational competency examination avoids these disadvantages. Two forms must be constructed, so that all test items and the tests are as similar as possible. The group takes the one form of test and, as soon as possible, the other form. The agreement between the group results on the two tests represents the coefficient of equivalence and stability. This method has certain advantages over the Test/Retest Method. A larger sampling of the occupation is possible and there is no need for concern about memory and intervening practice. However, the same disadvantages of availability of test population, limitation of examinees, and accessibility of facilities, make it less suited for occupational competency testing.

3. The Split Halves Method - Subdivision of a Test Into Equal Parts

a) When neither of the two procedures previously discussed is feasible, it may be practical to divide the test into halves. Usually, the odd numbers are contained in one test for which one set of scores is obtained for the group for one score. The even items are included in the second half for another score. Thus, two scores, which are reasonably equivalent, are obtained from a single test.

After the test is administered, two scores are obtained from the same test. The correlation coefficient, determined from these scores, reflects the reliability of the test. This estimate of reliability is often referred to as the split-half reliability coefficient or coefficient of equivalence. Since this value holds true for only one-half of the whole test, the reliability of the whole test must still be determined.

b) Calculating Reliability

The Spearman-Brown formula provides a means for estimating the whole test from its halves. The calculated reliability of the half test is substituted in the following equation:

$$\text{Reliability of whole test} = \frac{2\ (\text{reliability of half test})}{1 + (\text{reliability of half test})}$$

Two other split halves methods do not require computing the correlation coefficient. In the first method, only the standard deviation of each half test and of the total test are needed.

$$r = \frac{2\ (1 - So^2 + Se^2)}{St^2}$$

r = correlation coefficient
So = standard deviation of odd half
Se = standard deviation of even half
St = standard deviation of total test

The second formula yields the same results and requires only the standard deviation of the difference (Sd) between the half test scores and the standard deviation of the whole test.

$$r = 1 - \frac{Sd^2}{St^2}$$

If two parallel forms of the examination, written and performance, are constructed, the calculation of reliability is both simple and appropriate. The two forms of the examination are administered to the same group. Correlations are then calculated between the corresponding part-scores in the two forms, as well as the total scores correlations, the reliability coefficients of the test itself, and of the part-scores. This procedure is similar to the Gutman formula.

c) Other Methods of Estimating Internal Consistency

Kuder and Richardson do not require the splitting of the test into halves and rescoring and calculating a correlation coefficient.

$$r = \frac{n}{n-1} \cdot \frac{St^2 - (npq)}{St^2}$$

n = number of items in the test
St = standard deviation of the total test scores

$$p = \frac{\text{arithmetic mean of test scores}}{n} = \frac{Mt}{n}$$

$$q = 1 - p$$

This formula underestimates the internal consistency of a test if there is variation in difficulty among the items.

Another Kuder-Richardson formula, called Kuder-Richardson 20, requires the counting of the number of persons passing each item, dividing this by the number of persons taking the test, subtracting the resulting proportion (p) from (1) to obtain (q), getting (pq) from a table and adding the (pq's) for all items to get (Σpq). Since (p) is often valuable as part of an item analysis of a test in computing the difficulty of each item, the labor required in determining (p) for each item serves more than one purpose.

$$r = \frac{n}{n-1} \cdot \frac{St^2 - (\Sigma pq)}{St^2}$$

n = number of items on test
St = standard deviation of the total test scores
p = proportion of persons passing each item
q = 1 - p

145

d. Determining Reliability Through Analysis of Variance

For the performance examination, Hoyt's methodology of determining reliability through analysis of variance is suggested. The three sources of variance would be: (a) the difference between the mean scores of individuals in the group, (b) differences between the mean of the two forms, and (c) an error component.

$$\text{reliability} = \frac{A-B}{B}$$

A = mean square (difference between individuals)

B = mean square (error)

Another form of the same approach is the next formula:

$$re = \frac{\sigma t^2 - (\xi pq)}{2\sigma_t^2} + \sqrt{\frac{\xi r_c^2 + pq}{\sigma^2}\left\{\sigma^2_{2\sigma_t^2} - \xi pq\right\}}$$

The data required are:

1. The difficulties or percentages of correct items - P
2. The correlations between items and the total test score - rct
3. The standard deviation of the test σ_t

$$re = \frac{\sigma_t^2 - \xi pq}{2\sigma_t^2} + \sqrt{\frac{\xi r^2 + pq}{\sigma^2} + \frac{\sigma^2 - \xi pq}{2\sigma_t^2}}$$

This formula should be applied if an item analysis study has been made that provides the item test correlation coefficient and the difficulty value for each test item.

V. DIFFICULTY LEVEL OF TEST ITEMS

A. The Difficulty Index

The difficulty level of a test item provides some indication of the extent to which it does its job of measuring competency. The power to differentiate between levels of trade performance and knowledge is characteristic of a good test. For occupational competency test developers, this represents a difficult challenge because of the range of occupations. In some, the skills constitute the major portion of competence; in others, the more formal information and technical know-how determine proficiency. In still other trades or technical occupations, there is somewhat of a balance between manipulative skills and occupational information. A measure of difficulty level for test items may be obtained from the data for the combined high and low groups of scores achieved on a test.

The difficulty index is computed by dividing the number of candidates answering the item correctly by the total number in the high and low groups (high group usually within the upper 27% and the low group within the lower 27%). Example: If there are 16 candidates in the high group with 12 passing the item and 8 in the low group passing, the item difficulty index is:

Formula for Difficulty Index:

$$\text{Difficulty Index} = \frac{\text{Number of items answered correctly in high group}}{\text{Total number within the high group}} + \frac{\text{Number of items answered correctly in low group}}{\text{Total number within the low group}}$$

$$\text{Difficulty Index} = \frac{12 + 8}{16 + 16} = \frac{20}{32} = .625$$

The smallest possible value that can be obtained is zero and the largest 1. The larger the difficulty index, the easier the item.

B. Difficulty Distribution

The difficulty of an occupational competency test has an important bearing on its value. As a general rule, the average difficulty of the test item in a test should correspond to the average competency of craftsmen in the occupation, i.e., the items should be such that on the average about half the candidates will answer the test items correctly. However, in occupational competency testing, it is desirable to test for a range of skills and knowledge, the difficulty level of which must be spread over the range within which it is desired to differentiate. It should tend to cluster around 50% for the group.

A simpler version, adequate for most purposes is the following formula:

$$\text{Difficulty Index} = \frac{\xi x}{N} = \frac{\text{Total Right}}{\text{Total Number}}$$

This method does not require dividing the group into quartiles of high-medium-low, where the median consists of 50% of the candidates.

VI. DISCRIMINATION OF TEST ITEMS

A. Discrimination Index

While the difficulty level of an item determines, in part, the ability of the test item to discriminate between the proficiency of individuals, items of the same difficulty level do not always discriminate equally well. If 8 out of 16 individuals of a high group and 8 out of 16 in the low group pass, the item would have a difficulty index of .50, but would not discriminate between those who did well and those

who did poorly on the test as a whole. On the other hand, if 16 of the high group passed and none of the low group, its difficulty would, also, be .50. This might lead to the conclusion that the test had maximum discriminatory power.

A measure of discriminatory power may, also, be obtained on the basis of high and low groups of the wrong answers. The <u>discrimination index</u> may be computed by counting the wrong answers in the low group (W_l), and subtracting the wrong answers in the high group (W_h), and dividing by the number of candidates (N) in eigher the high or low group:

$$\text{Discrimination Index: } G = \frac{W_l - W_h}{N}$$

The discrimination for a group of 8 low and 8 high group scorers among a total group of 16 equals zero. If the low group was 16 and the high group zero, then the discrimination factor would be 1. When there are more wrong answers in the high group than in the low group, the discrimination factor becomes negative.

The range of values of the discrimination index depends upon the difficulty of the item. Items with difficulty levels of zero (0) or one (1) always have discrimination indices of zero. They cannot discriminate between craftsmen of low or high proficiency. The maximum discriminatory power drops as the item difficulty departs from .50. Since test makers strive for tests that have high discriminatory power, most test items should have difficulty levels between .40 and .60.

VII. LENGTH OF TESTS AND RELIABILITY

A. <u>Reliability Coefficient and Length of Test</u>

The reliability coefficient may be affected by the length of the test, the occupational proficiency range of the individuals to be tested, and the conditions under which the tests are administered. Other factors being equal, the reliability of a test is a function of its length; longer tests tend to be more reliable than shorter ones. The more samples taken of a certain area of skill, knowledge and/or behavior, the more reliable will be the appraisal of that area.

However, there is a limit beyond which a test of reliability is no longer practical. In using the formula:

$$rnn = \frac{nr}{1 + (n-1)r}$$

for a test of 50 items with a reliability coefficient of .80, the length of the test could be doubled to 100 items and raise the whole test reliability coefficient to .888, tripled to 150 items for a coefficient of .923 or quadrupled to 200 for a coefficient of .941. Note, in each instance, that the reliability coefficient increment from 50 to 100 items (.088) kept decreasing 100 to 150 (.035) to 200 items (.018). Eventually, this would level off.

B. Establishing Lengths of Tests

Should it become desirable to learn how much longer a test of known reliability should be made to achieve a particular reliability, the following Spearman-Brown formula may be used:

$$n = \frac{r_{nn}\ (1-r_{11})}{r_{11}\ (1-r_{nn})}$$

By properly substituting the values for r_{11} and r_{nn} in this equation, the multiple by which the number of test items should be increased can be obtained.

r_{nn} = .90(desired reliability), r_{11} = .70(test reliability)

$$n = \frac{.90\ (1-70)}{.70\ (1-90)} = \frac{(.90)\ \ (.30)}{(.70)\ \ (.10)} = 3.86$$

Example: A test of 50 test items has a reliability coefficient of .70. The question is, how many items should be included to establish a reliability coefficient of .90? Substituting values in the formula, the increase would be 3.86 x 50 = 193 test items.

VIII. RANGE OF ABILITY AND TEST SCORES

The reliability coefficient depends in part on the range of ability for the group for which the reliability coefficient is computed. A reliability coefficient for a test, based on a highly selected group of competitors, will be lower than the reliability based on an unselected, more variable group. Although the reliability coefficient may vary considerably with the range of ability, the standard error of the score may not be expected to do so. If it is assumed that the standard error of the scores is constant throughout the whole range of scores, it can be shown that the following relationship holds:

$$R = 1 - \frac{\sigma^2\ (1-r)}{\epsilon^2}$$

r= reliability coefficient for a group of test scores with standard deviation

R= reliability coefficient of the same test for a group of scores with a standard deviation

If a reliability coefficient of .80 has been obtained from a group of candidates for which the standard deviation if 10, then to find the estimated reliability for a group of subjects for which the standard deviation is 20,

$$R = 1 - \frac{100\ (1-.80)}{400} = .95$$

IX. RANGE OF ABILITY REGARDED AS STANDARD

No data is available for occupational competence tests which indicates a minimum coefficient of reliability. Among test specialists, the general opinion prevails that a value approximating .90 represents an acceptable standard.

X. ITEM ANALYSES OF TEST RESULTS

The total value of a test depends upon the sum total of the individual test items. For the test constructor, it is important to know:

A. What contribution each test item makes to the total test score,

B. How well the test scores discriminate in measuring proficiency among more or less competent individuals,

C. Do the options in the multiple-choice test items attract fairly well distributed responses,

D. Are there some choices which have been ignored by nearly all examinees, so that they might as well be excluded.

In the formulas that have been developed for establishing difficulty and discrimination data, test results for high, medium and low groups have been used. The high group constitutes 27% of the tested group, the low group 27%, and the remaining 46% represents the medium or average competence.

A typical item analysis is designed to provide the test constructor with information about the contribution of each test item to the total test score. Table XII shows an Item Analysis of Test Results. The rationale is that persons who score higher on the test, as a whole, should score higher on each test item than persons who do poorly on the test as a whole. Items 1 and 2 in Table XII reflect these expected differences, but items 3, 4 and 5 and 200 do not. The latter items should be reviewed by the test constructor to determine the cause of these differences.

Table XII. ITEM ANALYSIS OF TEST RESULTS BY CANDIDATES WITHIN SAME OCCUPATION

Item Number	Percent Passed Upper 27%	Percent Passed Lower 27%	Discrimination Index	Difficulty Index
1	.80	.40	.40	.40
2	.90	.20	.70	.45
3	.36	.50	-.40	.57
4	.66	.60	.06	.47
5	.98	.98	.00	.02
-	-	-	-	-
-	-	-	-	-
-	-	-	-	-
200	.40	.80	-.40	.40

The objective of item analysis is to improve the measuring instrument. It is possible to identify ambiguities, misleading words, totally irrelevent items and other errors or weaknesses.

XI. ANALYSIS OF ALTERNATE CHOICES

If the alternatives to the correct choice are obvious or are ignored by the examinees for other reasons, the test item loses its value to the overall test. The following examples demonstrate the steps in analyzing a multiple-choice test item:

Example 1: On a 100-point test, candidates made scores of 95-92-91, the other 22 candidates achieved scores ranging from 49 to 67. The measure of central position apt to be best for this group of 25 students is:

A - The Mode
B - The Median
C - The Mean
D - An Average of Mean and Mode

The responses to the choices are listed in Figure 16.

Figure 16. ANALYSIS OF EXAMPLE 1

Response	Choice				Difficulty	Discrimination
	A	B	C	D		
Upper 27%	1	17	-	-	0.72	.44
Lower 27%	5	9	3	1		
	6	26	3	1		

This item is relatively easy. Twenty-six students, out of 36 in the combined high and low group, marked the correct choice.

The difficulty index is calculated as $\frac{26}{36} = .72$

The discrimination index is calculated as $\frac{8}{18} = .44$

Three distractors attracted, at least, one member of the lower group. All are somewhat plausible to those who do not know the correct answer. This item might well be placed at the beginning of the test.

Example 2: In an oscillator, which uses a pie 20-electric effect for frequency control, energy is fed back to the grid circuit by

A. a tapped coil
B. inter-electrode capacity
C. a capacitor between plate and grid
D. inductive coupling
E. none of these

Figure 17. ANALYSIS OF EXAMPLE 2

Response	Choice					Difficulty	Discrimination
	A	B	C	D	E		
Upper 27%	12	3	-	2	1	.44	.44
Lower 27%	4	2	4	6	2		
	16	5	4	8	3		

This item appears to be more difficult than item 1 because fewer than one-half of the examinees, in the combined high and low group, marked the key answer (A). The difficulty factor is .44, and the discrimination index is .44. The maximum value of discrimination is $\frac{18-2}{18}$ or .89.
While this is low in discrimination, it distinguishes between high and low groups. Item 2 is closer to .50 in difficulty, and the distractors seem to be effective, particularly, for the low group.

Example 3: The relationship between the pulse repetition and the period is:

A. the higher the pulse rate the longer the period
B. the higher the pulse rate the shorter the period
C. the lower the pulse rate the shorter the period
D. none is correct

Figure 18 ANALYSIS OF EXAMPLE 3

Response	Choice				Difficulty	Discrimination
	A	B	C	D		
Upper 27%	-	3	-	15	.33	-.33
Lower 27%	-	9	-	9		
	-	12	-	24		

Obviously, there is a problem with this item. The item, in its present form, seems quite difficult. One-third of the group marked the item correctly. The proportion of correct answers is larger in the lower group. Therefore, it would seem easier, or it was confusing to the higher group. Options A and C were not selected by anyone.

The item was revised as follows:

Which statement describing the relationship between the pulse repetition and the period is correct?

A. the higher the pulse rate the longer the period
B. the higher the pulse rate the shorter the period
C. the lower the pulse rate the shorter the period
D. the relationship is not apparent

Figure 19 ANALYSIS OF REVISED EXAMPLE 3

Response	Choice				Difficulty	Discrimination
	A	B	C	D		
Upper 27%	-	18	-	-	.83	.33
Lower 27%	2	12	2	2		
	2	30	2	2		

The item has become relatively easy and discriminates reasonably well. The distractors are functioning.

These examples show one approach which can be used for test improvement. The statistical tools are of substantial help in test construction. They can point to weaknesses and point up the direction in which to proceed. Occupational competency testing involves many variables which are difficult to fit into a definite statistical pattern. Analysis and careful definition of the objectives, which lead to more objective criteria, are still an art. Eventually, they may become a science.

XII. TEST STANDARDS AND NORMS

A. Establishing Norms

A test may meet many or all the statistical requirements. The tabulated results show an acceptable spread among the individual scores, and the difficulty and discrimination coefficients fall into acceptable ranges. By themselves, these scores have little meaning. For, at what point, in the tabulated spread of scores, is the dividing line which separates the occupationally competent individual from the one whose skill and knowledge is not at a level acceptable for teaching others. In short, at what point does a candidate pass or fail?

A type of measure that assists in the interpretation of test scores is the norm. According to Adkins,

> "Norms are a measure of a specific function based on a specific sample that provides standards against which particular test scores can be interpreted."

The establishment of norms facilitates the setting of "pass" or "fail" levels. Such levels can be developed in measures of central tendencies, expressed in standard deviations, and in percentiles or stanines.

Assuming that the tests are truly representative of the competence of an occupation, a cutoff point at the 50th percentile would assure, at least, average proficiency. For more exacting standards, a cutoff point could be established at the 70th or 75th percentile. In such a situation, the candidate would be among the top quarter of the workers in a particular occupation with respect to occupational competence.

Such an arrangement would allow for variations among states and trade and industrial/technical teacher training institutions. For some occupations, certain subdivisions could be established with percentile ratings to provide greater flexibility in evaluating an individual's range of competence. To use occupational competency tests under varied conditions and requirements, norms must be established against which test results can be evaluated.

B. Types of Norms

Preference in achievement testing has been given to percentile norms. A number of other methods can be employed for a variety of purposes. It must be kept in mind, at all times, that raw scores achieved on a test are of little significance unless they indicate a degree of proficiency which can be related to some level of accepted proficiency in the occupation.

1. Computing Percentile Norms

Percentile norms are fairly easy to interpret. They express rank in terms of common units and can be computed by the following formula:

$$P = L + \frac{pN - f}{f_p} \cdot 1$$

L = bottom limit of the interval continuing the percentile

p = is the percentage of cases below the percentile

F = is the number of cases in all intervals below the interval containing the percentile

f_p = is the frequency within the interval containing the percentile

N = number of frequencies

1 = is the size of the class interval

Computing each of the 100 scores by formula is rather tedius. This computation can be obtained by computor or, more readily, by arithmetical or graphic interpolation from a smaller number of locations.

A more satisfactory method is to determine the percentile value of the top of each class interval by finding what percentage of the cases are covered by the cumulative frequency to the top of each interval and then to interpolate graphically for intermediate values. Table XIII illustrates the data needed for the second method.

The frequencies (f) in the second column are cumulative frequencies (column three) by successive addition from the lowest to the highest scores. The cumulative frequencies are converted to cumulative percentages by multiplying each entry by the reciprocal of N. These cumulative percentages (cumulative percent) are the percentiles corresponding to the tops of each score interval.

Table XIII. COMPUTATION OF PERCENTILE SCORES

Scores	(f) Frequencies	Cumulative Frequencies	Cumulative Percent
69-71	4	200	100
66-68	6	196	98
63-65	8	190	95
60-62	12	182	91
57-59	16	170	85
54-56	18	154	77
51-53	22	136	68
48-50	30	114	57
45-47	24	84	42
42-44	18	60	30
39-41	14	42	21
36-38	12	28	14
33-35	8	16	8
30-32	4	8	4
27-29	4	4	2
	200=N		

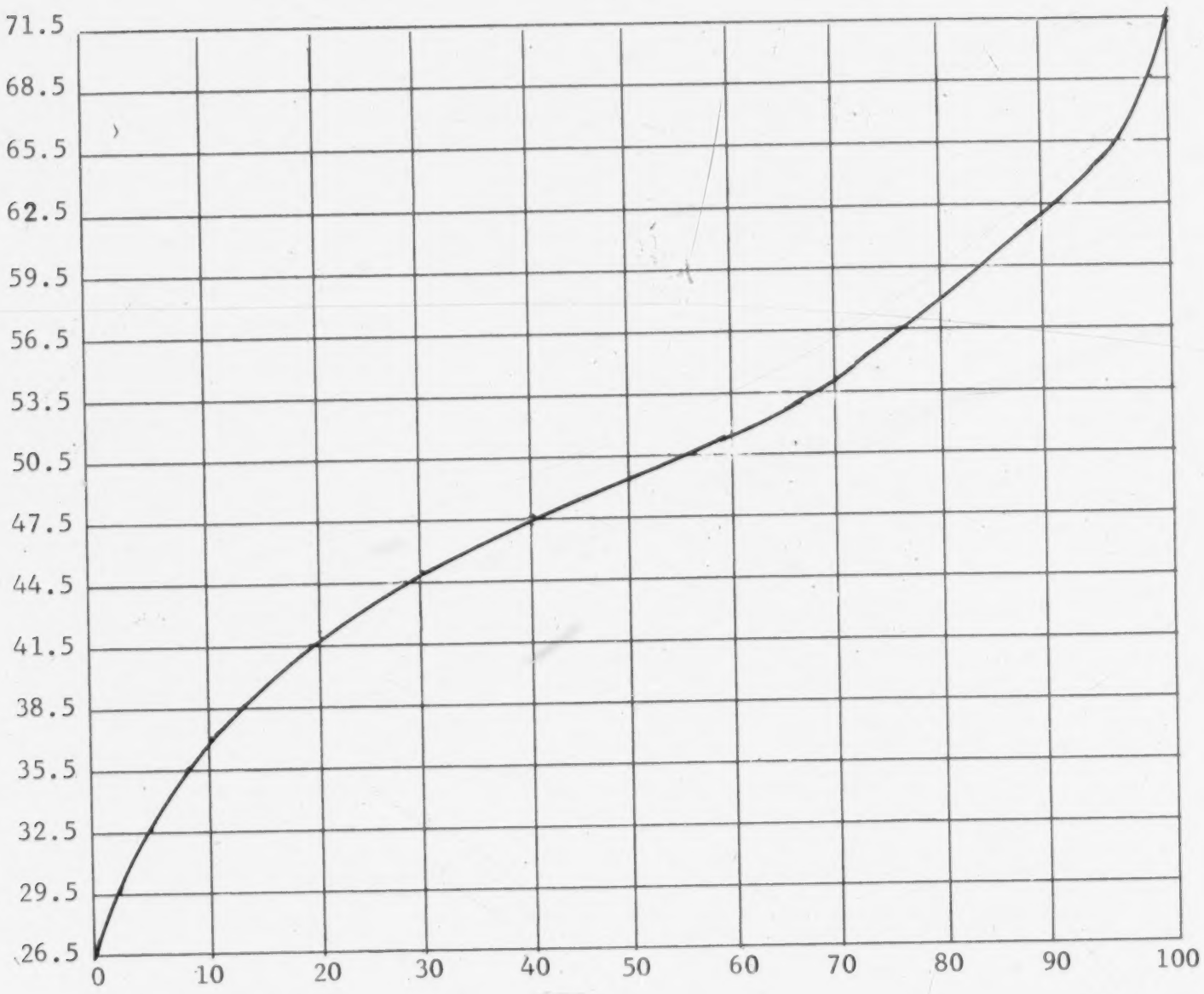
For greater convenience, accumulated percentiles may be plotted graphically on a percentile curve. Then, percentiles can be read directly without computation.

f = frequencies

Accumulated frequencies $4 + 4 + 8 + 12 = 28$

Percentiles For scores 36-38 $= \frac{\text{accumulated frequencies}}{\text{number of candidates}} \quad \frac{28}{200} = 14$

Figure 20 PERCENTILE GRAPH



To find the score for any desired percentile P, start from the P value on the horizontal axis, determine where a vertical line for this point cuts the vertical axis.

Percentile ranks provide very unequal units along the scale of performance on a test. The difference level in performance between scores at the 50th and 55th percentile is less than between scores at the 90th and 95th level in a bell-shaped distribution of raw scores, even though the differences in percentile rank are equal. Despite these disadvantages, percentile ranks are so easy to understand that they are used as norms for many standard tests.

2. Standard Scores

a. Z Score Defined

A standard score or Z score is defined as the deviation of a score from the arithmetic means in standard deviation units. The formula for calculating such scores is:

$$Z = \frac{x - m}{s}$$

Z = standard score
x = any raw score
m = arithmetic mean of raw score
s = standard deviation

If the mean of a distribution is 100 and its standard deviation 20, a candidate with a raw score of 120 receives a Z score of 1.00.

$$Z = \frac{120 - 100}{20} = 1.00$$

The advantage of the standard score, over the percentile rank for expressing the relative standing of an individual in his group, is that standard score units are equal over the entire scale so that small differences at one point on the scale mean exactly the same thing as an equal difference at some other point.

The chief value of standard scores is that they express scores in convenient and common units so that achievement on several tests, written and performance, can be understood and compared without reference to the original units.

$$Z = \frac{90 - 100}{20} = \frac{10}{20} = 0.50$$

Standard scores have the useful characteristics that their mean is always equal to 0 and their standard deviation equal to 1. All raw scores below the arithmetic mean (like the example with a raw score of 90) become negative. Several variations, like the derived score, have been introduced to eliminate this inconvenience.

$$D = 10 \frac{x - m}{s} + 50$$

A derived score (D) is simply defined as 10 times the Z score plus 50. The formula changes the mean (m) to 50 and the standard deviation (s) to 10.

b. Differences Between Standard Scores and Percentiles

A distinct difference exists between the standard score method and the percentile method. The standard scores are distributed more normally along the bell curve; the percentile scores tend to concentrate at the mean. Percentile scores cannot be averaged, the various standard scores can. It is, of course, an advantage to be able to average levels of performance when it is desired to average the results of several tests.

3. Normalized Standard Norms

The major features of percentile scores and standard scores can be combined into normalized standard scores by taking the following steps:

a) Convert raw scores into percentile scores

b) Convert percentile values into $\frac{x}{o}$ values. $\frac{x}{o}$ values are found in special tables of normal curve relationships which give values of (x) or distances from the mean which include (o) (specified proportions of the area or percentages of the cases).

Example: Determine the normalized standard score corresponding with a raw score of 47.

Percentile equivalent 80 is to be expressed in the standard score.

In a normal distribution, 50% of all cases fall below the mean score .-50 is subtracted from 80--- -50 = 30. Thirty percent of percentage cases fall between mean and percentile score of 80 in a normal distribution. The value of x ÷ o for an area of 30 is, approximately, .84. This value indicates the point on the base line of a normal curve which separates the lower 80 percent from the upper 20 percent.

Thus, .84 is the normalized standard score corresponding to a raw score of 47.

For raw scores corresponding to percentiles below 50, the standard score will be negative. A negative sign merely indicates that the distance referred to is below the mean.

In the determination of standards for a test, the conversion of one score to a corresponding score on another test must be kept in mind. Such conversions can be achieved by the use of special charts.

4. T-Scores

The T-scores serve the same purpose as the Z score. T-scores are based on the same principle, are always positive, and express larger units. Thus, T-scores eliminate the need for dealing with negative numbers and decimal fractions. A T-score is computed with the following formula:

$$T = 10 \frac{(x - m)}{s} + 50$$

x = any raw score
m = arithmetic mean of raw score
s = standard deviation of raw score

A T-score of 60 means a score of one standard deviation above the mean, a score of 70 two standard deviations above the mean, and so on. T-scores of 40 and 30, similarly, indicate scores at one and two standard deviations below the mean, respectively.

Chapter VIII

NATIONAL OCCUPATIONAL COMPETENCY TESTING PROGRAM: A NATIONAL URGENCY

I. THE NEED FOR A NATIONAL OCCUPATIONAL COMPETENCY TESTING PROGRAM

The key to the success of vocational education programs, at every level, is the occupationally competent teacher. The need for qualified tradesmen and technicians for teaching positions in vocational education programs is critical.

Determining occupational competency requires special personnel, is time consuming and expensive. Lack of such personnel and limited financial resources place a heavy burden upon those states which recognize the need for a sound testing program. In the past, personal interviews, review of employer's recommendations, oral examinations by experienced administrators and teacher educators have been used with some degree of success to select tradesmen and technicians for teaching positions. With the number of teachers required for the expanding vocational programs, objective information is needed which represents evidence as to an individual's competence.

Well designed and validated tests provide accurate and reliable information from which to evaluate an individual's occupational proficiency. The results of such tests can then be the basis on which to admit a candidate to an industrial teacher education program, provide the evidence required for certification, and enable institutions of higher learning to grant advanced standing for occupational proficiency in degree programs.

In this project, it was established that there are twenty centers, or state bureaus, which devote part of their effort to developing and administering competency tests. The survey further revealed a great deal of duplication of effort.

With the exception of a few states, limited evidence is available that the occupational competency tests have been subject to validation and reliability studies, or that standards or norms have been established; in short, that they are accurate and reliable instruments. It is, therefore, not surprising that state officials, as well as industrial teacher educators, have expressed serious concern for better and more accurate instruments.

In summary, it is obvious that objective, reliable and valid tests are more than ever necessary to measure occupational competence for institutional teaching, teaching within industry, within the military, and so forth. It is equally apparent that continued limited scattered efforts by the states can not produce the kinds of instruments needed.

The solutions to occupational competency test are to be found through a national coordinated effort under a broadly administered system, with built-in security and flexible controls that meet the needs of each state.

II. ADVANTAGE OF A NATIONAL TESTING PROGRAM

While a national effort must overcome a number of problems, the advantages well outweigh its disadvantages. Complementing the seven major and other additional functions defined in Volume I of this study are the following values for a National Testing Program:

A. Tests can be prepared more economically, thus conserving limited financial resources.

B. A pool of latest and most objective methods of test development can be applied to the preparation, administration, and evaluation of occupational competency tests.

C. Personnel now devoting part-time on a "hit-and-miss" basis can be relieved of test development responsibility to devote full-time to other program purposes.

D. Central data processing facilities can be more effectively utilized to evaluate tests and test results, and the techniques of statistical analysis may be more efficiently applied to create instruments that are accurate and reliable.

E. Tests can be more comprehensive in scope, yet provide greater flexibility to permit evaluation on various levels of competence within an occupation.

F. Tests, administered on a national basis, provide the information for the establishment of norms and standards essential for teaching in vocational programs.

G. Tests would be available to the states to use and interpret according to their own needs and certification requirements, and with built-in safeguards for state controls.

III. THE CONSORTIUM OF STATES AS A FACTOR IN RAISING STANDARDS AND RECOGNITION OF VOCATIONAL TEACHERS

A Consortium of States could establish an effective and efficient base for a National Occupational Competency Testing Center and Program. It could provide an administrative vehicle, develop administrative policies, set essential controls acceptable to each state, and lay out plans for implementation. Most important of all, by combining resources, evaluative instruments of high quality and acceptability could be developed which would:

A. Establish certain minimum required levels of competency;

B. Provide reciprocal arrangement, when all requirements for certification are met, for acceptance of individuals from one state to another;

C. Develop occupational competence levels among teachers on a broader national basis;

D. Encourage (through the granting of extensive content collegiate credits) vocational teachers to engage in studies in higher education leading to supervisory and administrator certificates;

E. Gain greater acceptance of occupational competency in institutions of higher learning; and

F. Establish a more professional approach toward certification and licensing of vocational teachers and gain greater professional recognition.

For the future vocational teacher, the successful passing of an occupational proficiency examination acknowledges that the individual tradesman or technician has the qualifications to function in one of the most important professions--that of a teacher.

APPENDICES OF RESOURCE MATERIALS

Appendix A. Selected Materials Arranged According to Sequence in Chapters

Appendix B. Selected References Arranged Alphabetically

APPENDIX A

I. SELECTED MATERIALS ARRANGED ACCORDING TO SEQUENCE IN CHAPTERS.

1. Cole, Lawrence F. and Bruce, William F.
Educational Psychology, pp-624-630 World Book Company, Yonkers on Hudson, N. Y. 1950

2. Fischer, George
Scale Books of the Reverend George Fischer in England (1864) with their "standard specimens" with which specimens of pupil work could be compared; also, the provocative article in The Forum in 1890's by its editor, Dr. J. M. Rice, who had measured thousands of American school children in spelling and other subjects by a series of tests.

3. Shimberg, Benjamin
The Performance Phase of Trade Competency Examination. Unpublished paper, Educational Testing Service, Princeton, N. J. 1966

4. Cole and Bruce
op. cit. pp 130

5. Chapman, J. Crosby
Trade Tests p. 24, Henry Holt & Company, New York, N. Y. 1921

6. United States Army
Performance Test Construction, United States Army Signal School; Educational Research Division, Office of Plans and Operations, Ft. Monmouth, N. J. 1966

7. Chapman
op. cit. p. 24

8. Thompson, L. E., Lawrence, D. C., Allardyce, A.
Interview Aids and Trade Questions for Employment Offices, Harper & Brothers, New York 1936

9. Chapman
op. cit. p. 24

10. Chapman
op. cit. p. 88

11. Stead, William and Chartle, Carol L.
Occupational Counselling Techniques Their Development and Application, American Book Company, New York, N. Y. 1940

12. United States Army, Performance Test Construction, Educational Research Division, Office of Plans and Operations, United States Arms Signal Corps School, Ft. Monmouth, N. J. p. 13 1966

13. Harns, C. W. (ed.)
The Encyclopedia of Educational Research, McMillan, 3rd ed., New York, N. Y. 1960

14. Krathwohl, D. R., Bloom, B. S., and Masia, B. B.,
Taxonomy of Educational Objectives. Affective Domain. David McKay, New York, N. Y. 1960

15. Simpson, Elizabeth Jane
"The Classification of Educational Objectives", Illinois Teacher of Home Economics 1966 10(4) 110-114

16. Moore, Maxine Ruth
The Proposed Taxonomy of the Perceptual Domain and Some Suggested Applications, Educational Testing Service, Princeton, N. J. 1960

17. Krathwohl, et. al.
op. cit.

18. Adkins
op. cit. p. 171

19. Lindeman
Educational Measurement, Scott, Foresman and Company, Glenview, Ill. pp..63 1967

20. Adkins
Illustrations derived from Adkins, op. cit. p. 33

21. The University of the State of New York
Improving the Class Room Test, The State Education Department, Bureau of Test Development, Albany, N. Y. 1964

22. Lindeman
op. cit. pp. 78

23. United States Navy Procedures for Constructing and Evaluating Achievement Examinations. Unpublished paper. Service School Command, United States Naval Training Center, Great Lakes, Ill. 1960
Adkins, op. cit. p. 41

24. Remmers, H. H., Gage, N. L. and Rummel, J. Francis
A practical Introduction to Measurement and Evaluation, Harper and Row, New York, N. Y. 1965

25. Michaels, W. J. and Karnes, M. R.
Measuring Educational Achievement, McGraw-Hill Book Company, Inc., New York, N. Y. p. 101 1950

26. United States Navy
op. cit.

164

27. Travers, M. W.
How to Make Achievement Tests, The Odyssey Press, New York, N. Y. p. 101 1950

28. Michaels
op. cit.

29. Lindquist, E. F. (Editor)
Educational Measurement, American Council on Education, Washington, D. C., George Banta Publishing Company, Menasha, Wisc. p. 214 1951

30. United States Army Performance Test Construction, Educational Research Division, Office of Plans and Operations, Ft. Monmouth, N. J. p. 17 1960

31. United States Army, op. cit.
Federal Aviation Administration, Designing Written Achievement Tests, Department of Transportation, Training Bulletin 103, FAA Academy 1965

32. Committee on Test Standards, AERA, and Committee on Test Standards, NCMUE, Technical Recommendation for Achievement Tests, National Education Association, Washington, D. C. 1955

Technical Recommendations for Psychological Tests and Diagnostic Techniques, Supplement to Psychological Bulletin 1954, No. 2, Pt. 2

33. Lindeman, Richard H.
Educational Measurement, Scott, Foreman and Company, Glenview, Ill. p. 37, 1967

34. Remmers, H. H.; Gage, N. L.; Rummel, J. Francis
A Practical Introduction to Measurement and Evaluation, Hopper and Row, New York, N. Y. p. 120 1965

35. Impellitere, Joseph T.
"Constructing Valid Occupational Competency Examinations", Unpublished paper in Feasibility of Providing Trade and Competency Examinations. Final Report. Rutgers, The State University, Graduate School of Education. Department of Vocational/Technical Education, New Brunswick, N. J. 1967

36. United States Army, Performance Testing, Enlisted Evaluation Center, Ft. Benjamin Harrison, Indianapolis, Ind. p. 9-1 1969

37. Gutman, L.
"A Basis for Analyzing Test - Re-Test Reliability", Psychometrica, 1945, 10: 255-282

38. Rulan, P. J.
"A Simplified Procedure for Determining the Reliability of a Test by Split-Halves", Harvard Education Review, 1934, 9: 99-103

39. Kuder, G. F. and Richardson, M. W.
The Theory of the Estimation of Test Reliability
Psychometrica, 1937 2:151-160

40. Remmers, et. al.
op. cit. p. 129

41. Hoyt, C. T.
Test Reliability Estimated by Analysis of Variance
Psychometrica, VI, 1941 pp. 153-160

42. Lindeman
op. cit. p. 89-91

43. Remmers, et. al.
op. cit. p. 131

44. Adkins, C. Dorothy; Primoff, S. Ernest; McAdaa, Harold L.; Bridges, Claude F.; Foree, Bertram
Construction and Analysis of Achievement Tests,
Superintendent of Documents, U. S Government Printing Office, Washington, D. C 1947, p. 145-160

45. Impellitere
op. cit.

46. Lindeman
op. cit. p. 92

47. Adkins, et. al.
op. cit. p. 137

48. Adkins, et. al.
op. cit. p. 137

49. Adkins, et. al.
op. cit. p. 140

50. Walker, Helen M.
Elementary Statistical Methods, Henry Holt & Company, New York, N. Y. 1943 p. 136-138

Adkins, et. al.
op. cit. p. 141

51. Walker
op. cit. p. 338-339

Kelley, Truman, L.
Statistical Method, MacMillan Company, New York 1923
p. 373-380

Guilford, Joy Paul
Psychometric Methods, McGraw-Hill, New York, N. Y. 1936
p. 530-534

52. Lindeman
op. cit. p. 133

Remmers
op. cit. p. 61

Adkins
op. cit. p. 142

53. United States Department of Labor
Occupational Licensing and the Supply of Non-Professional Manpower

Manpower Administration

Manpower Research Monograph #1

Superintendent of Documents, United States Government Printing Office, Washington, D. C. 1969 p. 1

APPENDIX B

II. SELECTED REFERENCES ARRANGED ALPHABETICALLY

Adkins, Dorothy C., Ernest S. Pirmoff and Harold L. McAdoo, Claude F. Bridges and Bertram Forerr, Construction and Analysis of Achievement Tests, Superintendent of Documents, U. S. Government Printing Office, Washington, D. C. 1947

Askren, William B. (Ed.), Symposium on Reliability of Human Performances in Work. Annual Convention of the American Psychological Association, U. S. A. F. - AMRL Technical Report, No. 67-68-45 pp. 1966

The A. P. A., AERA, NCME Joint Committee, Standards for Educational and Psychological Tests and Manuals, American Psychological Association, Washington, D. C. 1954

Anastasi, Anne, The Concept of Validity in the Interpretation of Test Scores. Education Psychological Measurement 1950-10-67-68

Anastasi, Anne, Psychological Testing, (2nd Edition), McMillan Company, New York, N. Y. 1965

Baldwin, Thomas S., The Development of Achievement Measures for Trade and Technical Education. Office of Education Grant-OEG2-7000517-0585, University of Illinois, Urbana, Ill. 1968

Bloom, B. S. (Ed.), Taxonomy of Educational Objectives: Handbook I, Cognitive Domain. Lougmans, Green and Company, N. Y. 1956

Bloomers, P and Lindquist, E. F., Elementary Statistical Methods in Psychology and Education. Houghton Hifflin, Boston, Mass. 1960

Buros, Oscar Krison, The Sixth Mental Measurement Yearbook. The Gryphon Press, Highland Park, N. J. 1961

Chapman, J. Crosby, Trade Tests, Henry Holt & Company, New York, N. Y. 1921

Cronbach, L. J., Essentials of Psychological Testing, Harper, N. Y. 1960

Digitek Corporation, Fairless Hills, Penn.

Fattu, N. A. and Medley, D. M., Rationale and Construction of the K-System Performance Test. An unpublished report of the Air Research and Development Command, Lowry Air Force Base, Denver. Colo. 1961

Gohem, Howard W., and Karuck, Samuel, Test Construction, U. S. Civil Service Commission, U. S. Government Printing Office, Washington D. C. 1951

Fleishman, Edwin A., Performance Assessment on an Emperically Derived Task Taxonomy. Human Factors Institute for Research, Washington, D. C. 1967

Furst, Edward J., Constructing Evaluation Instruments, Longmens, Green & Company, New York, N. Y. 1958

Fear, Richard A. and Byron, Jordon, Employee Evaluation Forms for Interviews, Psychological Corporation, New York, N. Y.

Green, Helen James, A Manual for Selected Occupational Tests for the Use in Public Employment Offices, Bulletins of the Employment Stabilization Research Institute, University of Minnesota, Vol. II, No. 3, July 1933 Minneapolis, Minn. 1953

I. B. M. Corporation, Binghamton, N. Y.

Krathwohl, David R., Bloom, Benjamin S. and Masia, Bertram B., Taxonomy of Educational Objectives Classification of Goals, Handbook II Affective Domain, David McKay Company, Inc., N. Y. 1964

Kynard, Alfred Tennyson, Criteria for Determining the Value of Work Experience for Teachers of Trade and Industrial Education, University of California, Los Angeles, Calif. 1960

Lindeman, Richard H., Educational Measurement, Scott, Foresman and Company, New York, N. Y. 1967

Lofgren, Paul V. W. and Allen, David, An Abridgement of Study of the Trade-Technical Teacher Proficiency Testing Program in California, Unpublished paper, Division of Vocational Education, University of California, Los Angeles, Calif. 1969

Measurement Research Center, Iowa City, Ia.

Moore, Maxine Ruth, A Proposed Taxonomy of the Perceptual Domain and Some Suggested Applications. Educational Testing Service, Princeton, N. J. 1958

Occupational Competency Tests, State Department of Education, Division of Vocational Education, Columbus, O. 1962

Panitz, Adolf and Olivo, C. Thomas, The State of the Art of Occupational Competency Testing, Rutgers University, Graduate School of Education, Department of Vocational/Technical Education, New Brunswick, N. J. 1970

Popham, James, Performance Tests of Instructors Competence for Trade & Industrial Education, Final Report, University of California, Los Angeles, Calif. 1968

Remmers, H. H., Gage, N. L. and Rummel, J. Francis, A Practical Introduction to Measurement and Evaluation, Harper and Row, New York, N. Y. 1965

Segal, David, State Testing and Evaluation Programs, Federal Security Agency, Circular No. 320, Washington, D. C. 1951

Shimberg, Benjamin, The Performance Phase of Trade Competency Examination, Unpublished paper, Educational Testing Service, Princeton, N. J. 1966

Siegel, Arthur, and Courtney, Douglas, Development of Practical Performance Measures, (Under contract with the Office of Naval Research) Institute for Research in Human Relations, Philadelphia, Pa. 1953

Stolurow, Lawrence M., The Taxonomy of Learning Task Characteristics, Technical Documentary Report AMRL-TDR-64-2, Behavioral Sciences Laboratory, Wright Patterson Air Force Base, Ohio 1964

Stuitt, Dewey (Ed.), Personnel Research and Test Development in the (U.S.) Bureau of Naval Personnel, Princeton University Press, Princeton, N. J. 1947

Super, D. E., Appraising Vocational Fitness by Means of Psychological Tests, Harper, New York, N. Y. 1949

Swanson, R. A., Auditory Automotive Mechanics Diagnostic Achievement Tests, (Unpublished doctoral dissertation) University of Illinois, Chicago, Ill. 1968

The Ohio State University, Instructions to Test Administrators, Instructions Materials Laboratory, Trade and Industrial Education Science, State Department of Education, Division of Vocational Education, Columbus, O. 1965

The University of the State of New York, Improving the Class Room Test, The State Education Department, Bureau of Test Development, Albany, N. Y. 1964

Thompson, Lorin Andrew and others, Interview Aids and Trade Questions for Employment Offices, Harper and Brothers, New York, N. Y. 1936

Thorndike, Robert L. and Hagen, Elizabeth, Measurement and Evaluation in Psychology and Education, John Wiley & Sons, New York, N. Y. 1961

Tinkelman, Sherman, Difficulty Prediction of Test Items, Bureau of Publications, Teachers College, Columbia University, New York, N. Y. 1947

U. S. Army Enlisted Evaluation Center, Performance Testing, Ft. Benjamin Harrison, Indianapolis, Ind. 1969

U. S. Department of Labor, Occupational Licensing and the Supply of Non-Professional Manpower, Manpower Administration, Manpower Research Monograph #1., Superintendent of Documents, U. S. Government Printing Office, Washington, D. C. 1969

U. S. Army Signal School, Performance Test Construction, Ft. Monmouth, N. J. 1960

U. S. Navy, Procedures for Constructing and Evaluating Achievement Examinations, Unpublished Paper, Service School Command, U. S. Naval Training Center, Great Lakes, Ill.

U. S. Office of Education, Teacher Competency in Trade and Industrial Education, U. S. Department of Health Education & Welfare, OE 84006, Washington, D. C. 1960

Wood, D. A., Test Construction, Charles Merril Books, Columbus, O. 1960

Wipple, James E., Baldwin, Robert D., Mager, F. and Vineberg, Robert, A Three Hour Performance Test to Evaluate Job Effectiveness of Army Radar Mechanics, Technical Report 38 Human Resources Research Office, Alexandria, Va. 1965

Walker, H. M. and Lew, J., Statistical Inference, Holt, Rinehart & Winston, New York, N. Y. 1953